OF MICHIGAN
DEC 1955
MATH. ECON.
LIBRARY

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

REINHOLD BAER
UNIVERSITY OF ILLINOIS

WEI-LIANG CHOW
THE JOHNS HOPKINS UNIVERSITY

ANDRÉ WEIL
UNIVERSITY OF CHICAGO

AUREL WINTNER
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

S. S. CHERN
C. CHEVALLEY
J. A. DIEUDONNÉ
A. M. GLEASON
HARISH-CHANDRA
P. HARTMAN
G. P. HOCHSCHILD
I. KAPLANSKY
E. R. KOLCHIN
W. S. MASSEY
D. C. SPENCER
A. D. WALLACE

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LXXVII, Number 4
OCTOBER, 1955

THE JOHNS HOPKINS PRESS
BALTIMORE 18, MARYLAND
U. S. A.

CONTENTS

PAGE

On the generators of an ideal, with an application to the geometry of numbers in unitary space U_2. By K. ROGERS, . . . 621
An outline of the theory of F_2 series. By S. VERBLUNSKY, . . 628
Isomorphism between H^p and L^p. By R. P. BOAS, JR., . . . 655
On finite groups with cyclic Sylow subgroups for all odd primes. By MICHIO SUZUKI, ... 657
Asymptotic integrations of ordinary non-linear differential equations. By PHILIP HARTMAN and AUREL WINTNER, . . . 692
Primitive alternative rings and semi-simplicity. By ERWIN KLEINFELD, 725
Appendix to the paper "Mean periodic functions I." By LEON EHRENPREIS, ... 731
On the regular solutions at a point of singularity of a system of non-linear differential equations. By ROBERT W. BASS, . . 734
Representations of semisimple Lie groups IV. By HARISH-CHANDRA, . 743
Invariants of finite groups generated by reflections. By CLAUDE CHEVALLEY, ... 778
On the arithmetic structure of a class of commutative semi-groups. By LEE M. SONNEBORN, ... 783
On a problem of Poincaré concerning Riccati's equation. By PHILIP HARTMAN and AUREL WINTNER, ... 791
Isomorphisms between projective unitary groups. By JOHN H. WALTER, 805
On the local embedding problems in the differential geometry of surfaces. By AUREL WINTNER, ... 845
On indefinite binary Riemannian metrics. By AUREL WINTNER, . . 853
On the Galois theory of differential fields. By E. R. KOLCHIN, . . 868
Engel rings and a result of Herstein and Kaplansky. By M. P. DRAZIN, 895
A characterization of plane continuous curves which are filled up by continuous collections of continuous curves. By MARY-ELIZABETH HAMSTROM, ... 914
Commutator subgroups of free groups. By MAURICE AUSLANDER and R. C. LYNDON, ... 929
Errata, ... 932

The AMERICAN JOURNAL OF MATHEMATICS appears four times yearly.

The subscription price of the JOURNAL is $8.50 in the U. S.; $8.75 in Canada; and $9.00 in other foreign countries. The price of single numbers is $2.50.

Manuscripts intended for publication in the JOURNAL should be sent to Professor AUREL WINTNER, The Johns Hopkins University, Baltimore 18, Md.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE 18, MARYLAND, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

ON THE GENERATORS OF AN IDEAL, WITH AN APPLICATION TO THE GEOMETRY OF NUMBERS IN UNITARY SPACE U_2.*

By K. Rogers.

1. In the course of joint work with H. P. F. Swinnerton-Dyer on the geometry of numbers in unitary space, U_n denoting n-dimensional vector-space over the field of complex numbers, the author proved some results on algebraic numbers. These are given below and applied to the general theory of lattices in U_2, with special reference to the problem of two linear forms with complex coefficients and variables integers of an algebraic number field.

Let k be the field of rationals, θ a number algebraic over k, and $k(\theta)$ the field obtained by adjunction of θ. 'Integer' will mean 'integer of $k(\theta)$' unless otherwise specified. Every ideal $\mathfrak{a}$ can be expressed as (p, q), where p, q are integers which we shall call a *generating pair*. Also, every ideal has a finite *basis* as module over the rational integers, and in particular there is a basis of two elements when θ is quadratic over k. We prove the following results:

THEOREM 1. *If $(p, q) = (r, s) = \mathfrak{a}$, then for any integer u of $k(\theta)$ there exists an integral 2×2 matrix A of determinant u, such that $(p, q)A = (r, s)$.*

When $k(\theta)$ is quadratic, we have in particular:

COROLLARY. *If (a, b) is a basis of $\mathfrak{a}$, then all generating pairs of $\mathfrak{a}$ are obtained from (a, b) by transformations with coefficients integral and determinant a unit.*

Calling this type of transformation 'integral unimodular,' we also prove:

THEOREM 2. *If $\mathfrak{M}(e)$ is the set of all 2×2 matrices A with integral elements and* Norm$(\det A) \leqq e$, *then there exists a finite subset $\{U_1, \cdots, U_h\}$ of $\mathfrak{M}(e)$ such that to any $A \in \mathfrak{M}(e)$ there corresponds a U_m and a unimodular integral matrix T such that $A = U_m T$.*

The referee pointed out that the case $u = 1$ of Theorem 1 is *Satz* 53 in Hilbert's *Zahlbericht* and is due to Hurwitz [1]. Also, Swinnerton-Dyer

* Received October 8, 1954; revised May 27, 1955.

and I have proved Theorem 2 for $n \times n$ matrices by considering lattices in U_n, but the proofs of both theorems given here are needed for the application to linear forms.

2. Proof of Theorem 1. Since $\mathfrak{a} = (p, q) = (r, s)$, there exist integers a, b, c, d such that $r = ap + bq$, $s = cp + dq$. Hence,

$$r = (a+f)p + (b-g)q, \qquad s = (c+f')p + (d-g')q,$$

where f, g, f', g' are integers such that

$$fp = gq, \qquad f'p = g'q. \tag{2.1}$$

Clearly, g and g' are any integers such that qg/p and qg'/p are also integers, that is, are any elements of the ideal $(p)/(q, p)$. It remains to choose solutions of (2.1) such that $(a+f)(d-g') - (b-g)(c+f') = u$. Denoting this determinant by Δ, setting $ad - bc = \Delta_0$, and noting that $fg' = f'g$, we have

$$\Delta = \Delta_0 - ag' + df + cg - bf',$$

and hence

$$\begin{aligned} p\Delta &= p\Delta_0 - apg' + dpf + cpg - bpf' \\ &= p\Delta_0 - apg' + dqg + cpg - bqg' \\ &= p\Delta_0 + gs - g'r. \end{aligned} \tag{2.2}$$

Now $gs - g'r$ takes any values in the ideal $(p)(r, s)/(q, p) = (p)$, and in particular the value $p(u - \Delta_0)$. With this value in (2.2), we have $\Delta = u$, as required.

3. Proof of Theorem 2. By the bound on $\mathrm{Norm}(\det A)$, the number of different ideals $(\det A)$ for $A \in \mathfrak{M}(e)$ is finite. We need therefore only prove the theorem for matrices of determinant E, say, where $\mathrm{Norm}(E) \leqq e$, since unit factors can be absorbed into the unimodular matrices. Let A denote such a matrix, and write

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Since $ad - bc = E$, the ideal $\mathfrak{a} = (a, b)$ must divide E and is therefore one of a finite number of ideals, finiteness depending on E and hence on e. For each of these ideals we can choose generating pairs, and then, by Theorem 1, we can multiply A on the right by a unimodular integral matrix such that

the new first row is one of a finite number. If it happens that the row-ideal is principal, then we can find a unimodular T such that

$$AT = \begin{pmatrix} a' & 0 \\ b' & c' \end{pmatrix},$$

in which case $a'c' = E$, and so apart from units a' and c' belong to a finite set of integers depending on e. Then we can reduce b' modulo c', this corresponding to multiplication on the right by a unimodular matrix, by substracting multiples of the second column from the first. In this way we should have found a unimodular integral matrix T such that AT is a matrix whose elements are integers out of a finite set depending only on e. Since T has an integral unimodular inverse, the theorem is proved in the particular case just considered. We now suppose that the ideal $\mathfrak{a}$ is non-principal, so that if

$$AT = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix},$$

where a' and b' belong to a finite set of integers, then $a'b'$ is not zero. We now consider what unimodular transformations can be applied so as to leave the first row invariant and to vary the element c' modulo some integer which is bounded in terms of e. We have

$$\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}\begin{pmatrix} 1+f_1 & f_2 \\ -g_1 & 1-g_2 \end{pmatrix} = \begin{pmatrix} a' & b' \\ c'+c'f_1-d'g_1 & \cdots \end{pmatrix}$$

where we require that

$$a'f_i = b'g_i \qquad (i = 1, 2). \tag{3.1}$$

We wish also to satisfy $\begin{vmatrix} 1+f_1 & f_2 \\ -g_1 & 1-g_2 \end{vmatrix} = 1$, or

$$f_1 = g_2. \tag{3.2}$$

We can certainly take $f_1 = g_2 = ua'b'$, where u is any integer, and then $f_2 = ub'^2$, $g_1 = ua'^2$. In this case

$$c'f_1 - d'g_1 = ua'(c'b' - a'd') = -a'uE,$$

so that the element c' can be varied modulo $-a'E$. In certain special cases, which are important in applications, we can vary c' modulo E. These are the cases when a' divides b'^2, when we can take the set of values

$$f_1 = g_2 = ub', \qquad g_1 = ua', \qquad f_2 = ub'^2/a',$$

which then give $c'f_1 - d'g_1 = u(c'b' - d'a') = -uE$. An example of the

second possibility is the ideal $(2, 1+i\sqrt{5})$ in the ring of integers of $k(i\sqrt{5})$. In any case, we have now found a unimodular integral matrix T_1 such that

$$AT_1 = \begin{pmatrix} a' & b' \\ c'' & d'' \end{pmatrix},$$

where now a', b' and c'' are bounded in terms of e. But since a' is not zero we can write $d'' = (E + b'c'')/a'$, and so d'' also has only a finite number of possible values. This completes the proof of Theorem 2.

4. Applications of these results. We require some definitions analogous to those of Mahler [3]. A function $f(\boldsymbol{x}) = f(x_1, \cdots, x_n)$ is called a *distance function* in U_n if

1) $f(\boldsymbol{x})$ is continuous and real-valued for complex x.

2) $f(\boldsymbol{x}) \geqq 0$ for all $\boldsymbol{x} \varepsilon U_n$ and $\neq 0$ at at least one point.

3) $f(t\boldsymbol{x}) = |t| f(\boldsymbol{x})$, for all $\boldsymbol{x} \varepsilon U_n$ and all complex t.

The set K of all points satisfying $f(\boldsymbol{x}) \leqq 1$, where $f(\boldsymbol{x})$ is a distance function, is called a *star-body*. A *lattice in* $k(\theta)$ is defined as the set Λ of all points $\boldsymbol{x} = A\boldsymbol{u}$, $(\det A \neq 0)$ where $\boldsymbol{u}$ runs through all $n \times 1$ column-vectors with elements integers of $k(\theta)$. Since $T\boldsymbol{u}$, for unimodular integral T, runs through all such column-vectors, the matrix AT also defines the same lattice. Since $|\det(AT)| = |\det A|$, this is a well-defined positive number which is called the determinant of the lattice, $d(\Lambda)$. A lattice Λ is said to be *admissible* for a star-body K if K contains no point of Λ in its interior other than the origin. Since the origin is an inner point of K, it is not difficult to show that there exists a positive number c, depending only on K, such that all lattices with $d(\Lambda) < c$ are inadmissible for K. We define the *critical determinant* $\Delta(K)$ as the lower bound of the determinants of admissible lattices, with the convention that $\Delta(K) = \infty$ if there is no admissible lattice. The preceding remark implies that for star-bodies $\Delta(K) > 0$. A lattice of determinant $\Delta(K)$ which is admissible for K is called a *critical* lattice of K. The central problem of the geometry of numbers is to find $\Delta(K)$ for various types of star-bodies and to find all possible critical lattices. Our object will be to prove that for star-bodies in U_2 with finite critical determinant, and for $k(\theta)$ an imaginary quadratic field over k, critical lattices always exist. For this we define a *bounded sequence of lattices* $\{\Lambda_r\}$ as a sequence for which there exist constants c_1, c_2 such that $d(\Lambda_r) \leqq c_1$ and $|\boldsymbol{x}| = \sqrt{\boldsymbol{x}^*\boldsymbol{x}} \geqq c_2$ for all $\boldsymbol{x} \varepsilon \Lambda_r$, $\boldsymbol{x} \neq 0$, $r = 1, 2, \cdots$.

LEMMA 1. *If $k(\theta)$ is complex quadratic, then every bounded infinite sequence of lattices in U_2 contains a convergent infinite subsequence.*

Proof. We begin by showing that Λ_r has independent points P_r, Q_r bounded in terms of c_1 and c_2, and for this we use a theorem of H. Weyl [6]. If $f(\boldsymbol{x})$ is a distance-function in U_n, let V denote the content of the star-body corresponding to that given by $f(\boldsymbol{x}) \leqq 1$ obtained by taking the real and imaginary parts of each co-ordinate as a co-ordinate in Euclidean $2n$-space. Let M_1 be the smallest positive number such that the region $f(\boldsymbol{x}) \leqq M_1$ has on its boundary a point of a given lattice Λ, say $\boldsymbol{x}^1$, and define M_k inductively as the least M such that $f(\boldsymbol{x}) \leqq M$ has on its boundary a point $\boldsymbol{x}^k$ of Λ which is independent of $\boldsymbol{x}^1, \cdots, \boldsymbol{x}^{k-1}$. By applying a result of Minkowski, Weyl showed that

$$M_1^2 \cdots M_n^2 V \leqq (2d^{\frac{1}{2}})^n \{d(\Lambda)\}^2,$$

where $-d$ is the discriminant of the quadratic field $k(\theta)$. In the present case we take

$$f(x) = \sqrt{\boldsymbol{x}^*\boldsymbol{x}} = \sqrt{(|x_1|^2 + |x_2|^2)},$$

so that V is the content of the 4-dimensional sphere of unit radius. Since $M_1 \geqq c_2$, we have

$$M_1^2 \leqq M_2^2 \leqq 4d\{d(\Lambda)\}^2 V^{-1} c_2^{-2},$$

and so that points $P_r = \boldsymbol{x}_r^1$, $Q_r = \boldsymbol{x}_r^2$ are of non-zero determinant and are bounded independently of r. Taking a matrix A_r which represents the lattice, we suppose that $\boldsymbol{x}_r^1 = A_r \boldsymbol{u}_r^1$, $\boldsymbol{x}_r^2 = A_r \boldsymbol{u}_r^2$, noting that if we take the matrix $A_r T^{-1}$ to represent the lattice, then the corresponding integer vectors are $T\boldsymbol{u}_r^1$ and $T\boldsymbol{u}_r^2$. Since

$$\det(\boldsymbol{x}_r^1, \boldsymbol{x}_r^2) = \det A_r \cdot \det(\boldsymbol{u}_r^1, \boldsymbol{u}_r^2),$$

and since $|\det A_r|$ has a lower bound independent of r, owing to the fact that Λ_r is admissible for the sphere of radius c_2, we have an upper bound for $|\det(\boldsymbol{u}_r^1, \boldsymbol{u}_r^2)|$ which is independent of r. By Theorem 2 applied to transposed matrices, we conclude that there exists a finite set of integral matrices $\{U_1, \cdots, U_m\}$ such that for each r there exists a unimodular integral matrix T_r with $T_r(\boldsymbol{u}_r^1, \boldsymbol{u}_r^2) = U_i$ for some i, $1 \leqq i \leqq m$. Thus we can choose a subsequence of $\{\Lambda_r\}$ such that for each lattice there is a basis for which the points P, Q have co-ordinates independent of r. We take this subsequence and rename it $\{\Lambda_r\}$ for convenience. Since the values of $d(\Lambda_r)$ are bounded, we can choose an infinite subsequence of $\{\Lambda_r\}$ such that $d(\Lambda_r)$ tends to some positive number, say Δ. Renaming this subsequence $\{\Lambda_r\}$, for convenience of notation, we now have a sequence of lattices $\{\Lambda_r\}$ and point-pairs P_r, Q_r, such

that $d(\Lambda_r) \to \Delta$, P_r and Q_r are bounded independently of r, and there is a matrix A_r representing Λ_r such that $(P_r, Q_r) = A_r U$, where U is an integral matrix independent of r. Since P_r and Q_r are bounded independently of r, we can choose a subsequence of $\{\Lambda_r\}$, which we rename once more, such that all the above conditions are satisfied and also $P_r \to P$, $Q_r \to Q$, as $r \to \infty$. Then, since U is fixed and non-singular, the fact that $A_r U$ tends to a limit as $r \to \infty$ implies that there exists a matrix A of determinant of modulus Δ, such that $A_r \to A$ as $r \to \infty$. Hence this subsequence of the original $\{\Lambda_r\}$ converges to the lattice Λ corresponding to A in the sense that there exists a basis of Λ (represented by the columns of A) to which a basis of a lattice of the sequence converges as $r \to \infty$. The lemma is proved.

THEOREM 3. *With respect to the complex quadratic field $k(\theta)$, every star-body in U_2 of "finite type" (that is, with finite $\Delta(K)$) possesses at least one critical lattice.*

Proof. By the definition of $\Delta(K)$, there exists a sequence $\{\Lambda_r\}$ of lattices admissible for K such that $d(\Lambda_r) \to \Delta(K)$ as $r \to \infty$. By discarding a finite number of these, we can suppose that $d(\Lambda_r) \leqq 2\Delta(K)$. As we remarked earlier, there exists a sphere round the origin which lies entirely in K; hence this and the previous inequality imply that the sequence $\{\Lambda_r\}$ is bounded. By the lemma there is a subsequence, which we rename $\{\Lambda_r\}$, which converges to a lattice Λ of determinant $\Delta(K)$. To prove that Λ is critical for K it remains only to prove that Λ is admissible for K. Taking bases (P_1, P_2) and (P_{r1}, P_{r2}) such that $(P_{r1}, P_{r2}) \to (P_1, P_2)$ as $r \to \infty$, suppose the point $Q = u_1 P_1 + u_2 P_2 \neq 0$ of Λ lies in the interior of K. Then the point $Q_r = u_1 P_{r1} + u_2 P_{r2} \neq 0$ of Λ_r tends to $u_1 P_1 + u_2 P_2$ as $r \to \infty$, and hence, since the distance function of K is continuous, the point Q_r lies in the interior of K for large enough r. This contradicts the fact that Λ_r is K-admissible, and hence Λ must be a critical lattice of K.

5. The above theorem has practical application, quite apart from its interest as a pure existence theorem. A lattice Λ is said to be *extremal* for K if in a sufficiently small 'neighbourhood of Λ' (defined in an obvious way) there are no K-admissible lattices Λ' such that $d(\Lambda') < d(\Lambda)$; clearly a critical lattice is extremal. Thus, if we have *necessary* conditions for an extremal lattice, and if these are satisfied by only a finite number of lattices, then since critical lattices exist we know that a lattice of smallest determinant is a critical lattice. This has been applied by the author to the problem of maximising the function

$$\min_{u,v} \{\max(|\alpha u + \beta v|, |\gamma u + \delta v|\}$$

of complex numbers $\alpha, \beta, \gamma, \delta$ such that $\alpha\delta - \beta\gamma = \Delta \neq 0$, the above minimum being taken over all integers u, v of $k(\theta)$ which are not both zero. With the help of an important lemma of Swinnerton-Dyer, the author found simple proofs for the cases $k(i)$ and $k(\rho)$, first considered by Minkowski [4]. We have now found the best possible result for $k(i\sqrt{5})$, when new problems arise from the class-number being 2. One has two columns of integers, the rows representing lattice points and generating either (1) or $(2, 1 + i\sqrt{5})$. The determinants are bounded, and we wish to reduce the array by unimodular right-multiplication to a finite number of possible cases. The proof of Theorem 2 shows this is trivial if a row-ideal is (1), and otherwise one row can be taken as $(2, 1 + i\sqrt{5})$, while the first element of the next row can be reduced modulo the determinant E. The details of this result will be published in a joint paper with Swinnerton-Dyer. This will develop general theory in U_n and will include a complete analogue of results of Korkine and Zolotareff [2] and Swinnerton-Dyer [5].

In conclusion, I thank the referee for simplifying the proof of Theorem 1 and for the reference to Hurwitz' paper.

THE GRADUATE COLLEGE,
PRINCETON, N. J.

REFERENCES.

[1] A. Hurwitz, "Die unimodularen Substitutionen in einem algebraischen Zahlenkörper," *Mathematische Werke* (Verlag Birkhäuser, 1933), Bd. II, 244-268.

[2] A. Korkine and G. Zolotareff, "Sur les formes quadratiques positives," *Mathematische Annalen*, vol. 11 (1877), pp. 242-292.

[3] K. Mahler, "On lattice-points in n-dimensional star-bodies," *Proc. Royal Soc.* A, vol. 187 (1946), pp. 151-187.

[4] H. Minkowski, *Diophantische Approximationen* (Leipzig, B. G. Teubner, 1907), Chapter 6.

[5] H. P. F. Swinnerton-Dyer, "Extremal lattices of convex bodies," *Proc. Camb. Phil. Soc.*, vol. 49 (1953), p. 161.

[6] H. Weyl, "Theory of reduction for arithmetical equivalence," *Transactions of the American Mathematical Society*, vol. 48 (1940), pp. 126-164 (particularly Theorem 1*, p. 140).

AN OUTLINE OF THE THEORY OF F_2 SERIES.*

By S. VERBLUNSKY.

1. *Introduction.* Let $A(z)$, $B(z)$ be relatively prime polynomials in z, with real coefficients, such that the equation

$$A \sin \pi z - B \cos \pi z = 0 \tag{1}$$

has all its roots real and simple. Then any function $f(x)$ which is L-integrable in an interval of length 2π, say $(\alpha, \alpha + 2\pi)$, has an expansion of "Fourier" type, in terms of cosines and sines of $\mu_k x$. Here, $\mu_1 < \mu_2 < \cdots$ are the positive roots of (1), $\mu_{-1} > \mu_{-2} > \cdots$ are the negative roots, and, if 0 is a root, then $\mu_0 = 0$. ([3], Theorem 1.) To fix the ideas, we shall consider the case in which the roots are symmetrical with respect to the origin. It is known that this happens if and only if one of A, B is odd, and the other is even. ([3], Section 5.) The "Fourier" expansion can then be written

$$\tfrac{1}{2}a_0 + \sum_{k=1}^{\infty} a_k \cos \mu_k x + b_k \sin \mu_k x, \tag{2}$$

where $a_k - ib_k = 2\lambda_k \int_{\alpha}^{\alpha+2\pi} f(x) e^{-i\mu_k x}\, dx$, $(k = 1, 2, \cdots)$, and, if 0 is a root of (1), $a_0 = 2\lambda_0 \int_{\alpha}^{\alpha+2\pi} f(x)\, dx$, while, if 0 is not a root of (1), then $a_0 = 0$. The numbers λ_k are the residues, at $z = \mu_k$, of

$$Q(z) = \tfrac{1}{2}(A \cos \pi z + B \sin \pi z)/(A \sin \pi z - B \cos \pi z).$$

The object of this paper is to investigate the series (2) when the even one of A, B is of degree 2, and the odd one is of degree 1. If A is the even one, we shall call (2) the F_2A series of $f(x)$ in $(\alpha, \alpha + 2\pi)$. If B is the even one, we shall call (2) the F_2B series of $f(x)$ in $(\alpha, \alpha + 2\pi)$. It is no loss of generality to suppose that the leading coefficient of the even one is unity. Thus, for F_2A series we can write $A = z^2 + a$, $B = bz$, and for F_2B series, $A = az$, $B = z^2 + b$, and in both cases, $ab \neq 0$, since A, B are relatively prime.

The properties of an F_2A series depend on whether $b^2 + 4a$ is positive, zero, or negative. In this paper, we consider only the first case. Similarly

* Received January 18, 1955.

for F_2B series, we consider only the case in which a^2+4b is positive. On examining Theorems 3 and 4 of [3], we can make the following deductions. The points (a, b) such that, (i) $b^2+4a>0$ and (ii) such that the equation

$$(z^2+a)\sin\pi z - bz\cos\pi z = 0$$

has all its roots real and simple, are the points below the curve

$$y=-x,\ x\geqq 0; \qquad y=-2\,|\,x\,|^{\frac{1}{2}},\ x<0.$$

The points (a, b) such that, (i) $a^2+4b>0$ and (ii) such that the equation $az\sin\pi z-(z^2+b)\cos\pi z=0$ has all its roots real and simple, are not so easily described. However, they include all the points (a, b) with $a<0$, $b>0$, $b^2+4a>0$.

We now mention some of the results proved below, and, for brevity, we refer only to F_2A series.

The coefficients of an F_2A series in $(\alpha, \alpha+2\pi)$ are connected by two linear equations (Theorem 1). It is convenient to have a name for the series

$$\sum_2^\infty a_k\cos\mu_k x + b_k\sin\mu_k x, \tag{3}$$

derived from the F_2A series (2) by omitting the initial terms. We call (3) a "restricted F_2A series" (RF_2A series) in $(\alpha, \alpha+2\pi)$. Given an RF_2A series (3) in $(\alpha, \alpha+2\pi)$, the "missing" terms, viz. $\frac{1}{2}a_0+a_1\cos\mu_1 x+b_1\sin\mu_1 x$ are, by Theorem 5, determined to within an arbitrary additive multiple of

$$-(a/b)\cdot\sin\pi\mu_1/\mu_1+\cos\mu_1(x-\alpha-\pi).$$

An RF_2A series in $(\alpha, \alpha+2\pi)$ is an RF_2A series in $(\beta, \beta+2\pi)$ for all β (Theorem 4). An RF_2A series is summable $(C, 1)$ almost everywhere, and is the Fourier series of its Cesarò sum (in the sense of the theory of almost periodic functions), by Theorem 7. Theorems of the Young-Hausdorff type hold for F_2A series (Theorem 8). In enunciating such theorems, we must take account of the fact that there are two linear equations connecting the coefficients of an F_2A series. As an example, we quote the following special case of Theorem 8(a): Let c_ν, $\nu=2, 3, \cdots$, be a sequence of complex numbers such that $\sum|\,c_\nu\,|^p<\infty$, where $1<p\leqq 2$. Let $1/p+1/q=1$. Then there is precisely one real function $f\,\varepsilon\,L^q(\alpha, \alpha+2\pi)$ which satisfies the conditions

$$0=\lambda_0\int_\alpha^{\alpha+2\pi} f(t)\,dt, \qquad c_\nu=\lambda_\nu\int_\alpha^{\alpha+2\pi} f(t)\,e^{-i\mu_\nu t}dt. \qquad (\nu=2, 3, \cdots)$$

For this function,

$$\left(\int_\alpha^{\alpha+2\pi}|\,f\,|^q\right)^{1/q}\leqq\left(2\sum_1^\infty|\,c_\nu\,|^p/\lambda_\nu{}^{p-1}\right)^{1/p},\ \text{where } c_1=\lambda_1\int_\alpha^{\alpha+2\pi} f(t)\,e^{-i\mu_1 t}dt.$$

Finally, as an example of a uniqueness theorem, we have: If

$$\sum_1^\infty (a_k \cos \mu_k x + b_k \sin \mu_k x) \tag{4}$$

converges to zero for all x in the open interval $(-\pi, \pi)$, then (4) is obtained by termwise differentiation of the F_2A series of a constant in $(-\pi, \pi)$ (Theorem 9).

Sections 2-11. F_2A Series.

2. As mentioned in Section 1, we may suppose that

$$A = z^2 + a, \qquad B = bz, \qquad ab \neq 0. \tag{5}$$

Writing $A_k = A(\mu_k)$, etc., we have

$$\lambda_k = \tfrac{1}{2} \frac{A_k \cos \pi\mu_k + B_k \sin \pi\mu_k}{\pi(A_k \cos \pi\mu_k + B_k \sin \pi\mu_k) + A_k' \sin \pi\mu_k - B_k' \cos \pi\mu_k}.$$

But

$$A_k \sin \pi\mu_k = B_k \cos \pi\mu_k. \tag{6}$$

Hence

$$1/(2\lambda_k) = \pi + b(\mu_k^2 - a)/[(\mu_k^2 + a)^2 + \mu_k^2 b^2], \tag{7}$$

and, in particular,

$$1/(2\lambda_0) = \pi - b/a. \tag{8}$$

Let f be a real function of class $L(\alpha, \alpha + 2\pi)$. Write

$$c_k = \lambda_k \int_\alpha^{\alpha+2\pi} f(t) e^{-i\mu_k t} dt, \qquad (k = 0, \pm 1, \cdots),$$

so that $c_{-k} = \bar{c}_k$. Then the F_2A series of $f(x)$ in $(\alpha, \alpha + 2\pi)$ can be written $\sum c_k e^{i\mu_k x}$. The relation between this series and the function f in the interval $(\alpha, \alpha + 2\pi)$, will be denoted by

$$f(x) \sim \sum c_k e^{i\mu_k x}, \qquad (\alpha, \alpha + 2\pi). \tag{9}$$

If δ is a real number, then

$$f(x - \delta) \sim \sum c_k e^{-i\mu_k \delta} e^{i\mu_k x}, \qquad (\alpha + \delta, \alpha + \delta + 2\pi), \tag{10}$$

and

$$f(-x) \sim \sum c_{-k} e^{i\mu_k x}, \qquad (-\alpha - 2\pi, -\alpha). \tag{11}$$

3. As regards convergence, F_2A series behave like Fourier series. This is a special case of a known theorem ([3], Theorem 1), and will be enunciated as a lemma.

LEMMA 1. *If $s_n(x)$ denotes the n-th partial sum of* (9), *viz.,*

$$s_n(x) = \sum_{-n}^{n} c_k e^{i\mu_k x},$$

and $\sigma_n(x)$ denotes the n-th partial sum of the Fourier series of $f(x)$, viz.,

$$\sigma_n(x) = \sum_{-n}^{n} \frac{1}{2\pi} \int_{\alpha}^{\alpha+2\pi} f(t) e^{-ik(t-x)} dt,$$

then $s_n(x) - \sigma_n(x) \to 0$ uniformly in $(\alpha, \alpha + 2\pi)$ as $n \to \infty$. In particular, $s_n(\alpha + 2\pi) - s_n(\alpha) \to 0$ as $n \to \infty$.

In virtue of the last clause of Lemma 1,

$$\sum' c_k e^{i\mu_k \alpha}(e^{2\pi i \mu_k} - 1) = 0, \tag{12}$$

where here, and in what follows, the accent denotes that the term with $k = 0$ is absent.

The equation (1) can be wirtten in the form $e^{2\pi i z} = (A + iB)/(A - iB)$. We have $A - iB = z^2 - ibz + a = (z - im_1)(z - im_2)$. To keep our arguments as simple as possible, we shall suppose that $b^2 + 4a > 0$. Thus, m_1, m_2 are real and different. Then

$$e^{2\pi i \mu_k} - 1 = 2iB_k/(A_k - iB_k) = 2ib\mu_k/[(\mu_k - im_1)(\mu_k - im_2)]. \tag{13}$$

We can now prove that the numbers c_k in (9), which we may call the F_2A constants of f in $(\alpha, \alpha + 2\pi)$, are connected by two linear relations, independent of f.

THEOREM 1. *For any F_2A series* (9), *we have*

$$\sum' c_k e^{i\mu_k \alpha} \mu_k/[(\mu_k - im_1)(\mu_k - im_2)] = 0, \tag{14}$$

and

$$c_0/a + \sum' c_k e^{i\mu_k \alpha}/[(\mu_k - im_1)(\mu_k - im_2)] = 0, \tag{15}$$

or, what is equivalent, the two equations

$$\sum c_k e^{i\mu_k \alpha}/(\mu_k - im_1) = 0, \qquad \sum c_k e^{i\mu_k \alpha}/(\mu_k - im_2) = 0. \tag{16}$$

Proof. Equation (14) follows from (12) on substituting the value of $e^{2\pi i \mu_k} - 1$ from (13). To obtain (15), consider the F_2A series of the function defined as equal to 1 in $(\alpha, \alpha + 2\pi)$. If

$$1 \sim \sum c_k^* e^{i\mu_k x}, \qquad (\alpha, \alpha + 2\pi), \tag{17}$$

then

$$c_0^* = 2\pi\lambda_0, \quad c_k^* = (\lambda_k/-i\mu_k)\, e^{-i\mu_k\alpha}(e^{-2\pi i\mu_k} - 1), \qquad (k \neq 0). \tag{18}$$

On account of the uniform equiconvergence of an F_2A series and a Fourier series, we can multiply (17) by $f(x)$ and integrate term by term. We thus obtain

$$c_0/\lambda_0 = 2\pi c_0 + \sum' (c_{-k}/-i\mu_k)\, e^{-i\mu_k\alpha}(e^{-2\pi i\mu_k} - 1).$$

By (8) and (13), this becomes

$$2b/a \cdot c_0 + \sum' c_{-k}\, e^{-i\mu_k\alpha}\, 2b/[(\mu_k + im_1)(\mu_k + im_2)] = 0.$$

Divide by $2b$, and take conjugates. We obtain (15).

If we multiply (15) by $-im_2$ and add to (14), we obtain

$$-c_0 im_2/a + \sum' c_k\, e^{i\mu_k\alpha}/(\mu_k - im_1) = 0.$$

But $a = -m_1 m_2$. Hence the first term is $c_0\, e^{i\mu_0\alpha}/(\mu_0 - im_1)$. This proves the first equation (16). The second equation is proved similarly. By reversing the argument, we see that equations (16) imply (14) and (15).

4. In order to decide when the trigonometric polynomial

$$\sum_{|k| \leqq n} c_k e^{-i\mu_k x}, \qquad (c_{-k} = \bar{c}_k),$$

is an F_2A series in $(\alpha, \alpha + 2\pi)$, we need to examine two special cases. We first observe that if (n, ν) are positive integers, then by (6),

$$\int_{-\pi}^{\pi} \cos\mu_\nu x \cos\mu_n x\, dx = 2a/b \cdot \sin\pi\mu_\nu/\mu_\nu \cdot \sin\pi\mu_n/\mu_n, \qquad (\nu \neq n), \tag{19}$$

and

$$\int_{-\pi}^{\pi} \sin\mu_\nu x \sin\mu_n x\, dx = -2/b \cdot \sin\pi\mu_\nu \sin\pi\mu_n, \qquad (\nu \neq n). \tag{20}$$

LEMMA 2. (i) *If ν is a positive integer, then*

$$-(a/b) \cdot \sin\pi\mu_\nu/\mu_\nu + \cos\mu_\nu x \tag{21}$$

is its own F_2A series in $(-\pi, \pi)$.

(ii) *If $\nu, > 1$, is a positive integer, then*

$$-\sin\mu_1 x/\sin\mu_1\pi + \sin\mu_\nu x/\sin\mu_\nu\pi \tag{22}$$

is it own F_2A series in $(-\pi, \pi)$.

Proof. (i) The F_2A series of the even function (21) can be written $\frac{1}{2}a_0 + \sum_1^\infty a_n \cos \mu_n x$. If $n \neq 1, \nu$, then

$$a_n = 2\lambda_n \int_{-\pi}^{\pi} (-(a/b) \cdot \sin \pi\mu_\nu/\mu_\nu + \cos \mu_\nu x) \cos \mu_n x \, dx = 0,$$

by (19). Thus, the F_2A series of (21) is $\frac{1}{2}a_0 + a_\nu \cos \mu_\nu x$, and this converges to (21) in $(-\pi, \pi)$. Put $x = \pi/2\mu_\nu$. Then $\frac{1}{2}a_0 = -(a/b) \cdot \sin \pi\mu_\nu/\mu_\nu$. Hence also $a_\nu = 1$.

(ii) The proof is similar.

We can now solve the problem mentioned at the beginning of this section.

THEOREM 2. *In order that*

$$\sum_{|k| \leqq n} c_k e^{i\mu_k x}, \qquad (c_{-k} = \bar{c}_k), \tag{23}$$

shall be an F_2A series in $(\alpha, \alpha + 2\pi)$, it is necessary and sufficient that

$$\sum_{|k| \leqq n}' c_k e^{i\mu_k \alpha}(e^{2\pi i\mu_k} - 1) = 0, \tag{24}$$

and that

$$c_0/a + \sum_{|k| \leqq n}' c_k e^{i\mu_k \alpha}(e^{2\pi i\mu_k} - 1)/(2ib\mu_k) = 0; \tag{25}$$

or, what is equivalent, that

$$\sum_{|k| \leqq n} c_k e^{i\mu_k \alpha}/(\mu_k - im_1) = 0, \qquad \sum_{|k| \leqq n} c_k e^{i\mu_k \alpha}/(\mu_k - im_2) = 0. \tag{26}$$

Proof. That the conditions are necessary, follows from Theorem 1 on using (13). We further observe, on recalling the argument at the end of Section 3, that (24) and (25) are equivalent to (26).

To prove that the conditions (24) and (25) are sufficient, we may suppose that $\alpha = -\pi$. By Lemma 2, we can express (23) as the sum of an F_2A series in $(-\pi, \pi)$, plus an expression of the form

$$\gamma_0 + \gamma_1 e^{i\mu_1 x} + \gamma_{-1} e^{-i\mu_1 x}, \qquad (\gamma_{-1} = \bar{\gamma}_1), \tag{27}$$

where the γ's satisfy the conditions of the theorem, viz.

$$\sum_{k=\pm 1} \gamma_k e^{-i\pi\mu_k}(e^{2\pi i\mu_k} - 1) = 0, \tag{28}$$

and

$$\gamma_0/a + \sum_{k=\pm 1} \gamma_k e^{-i\pi\mu_k}(e^{2\pi i\mu_k} - 1)/(2ib\mu_k) = 0. \tag{29}$$

Now equation (28) is $(\gamma_1 - \bar{\gamma}_1)2i \sin \pi\mu_1 = 0$, and $\sin \pi\mu_1 \neq 0$ by (6). Thus γ_1 is real, and (29) can be written

$$\gamma_0/a + 2\gamma_1 \sin \pi\mu_1/b\mu_1 = 0. \tag{30}$$

Hence (27) is $2\gamma_1(-(a/b) \cdot \sin \pi\mu_1/\mu_1 + \cos \mu_1 x)$, and, by Lemma 2, is an F_2A series in $(-\pi, \pi)$. This completes the proof.

5. The F_2A series of $f(x)$ in $(\alpha, \alpha + 2\pi)$, being uniformly equiconvergent with the Fourier series of $f(x)$, is summable (C, 1) almost everywhere in the interval to $f(x)$. The first step in the investigation of the sum (C, 1) of the series outside the interval $(\alpha, \alpha + 2\pi)$ is to determine the value of the sum (C, 1) in the interval $(\alpha + 2\pi, \alpha + 4\pi)$.

THEOREM 3. *If*

$$f(x) \sim \sum c_k e^{i\mu_k x}, \qquad (\alpha, \alpha + 2\pi), \tag{31}$$

then $\sum c_k e^{i\mu_k x}(e^{2\pi i\mu_k} - 1)$ *is uniformly convergent for* $\alpha \leqq x \leqq \alpha + 2\pi$, *and its sum is*

$$-[2b/(m_1 - m_2)][m_1 e^{-m_1 x} \int_\alpha^x f(t)e^{m_1 t}\,dt - m_2 e^{-m_2 x} \int_\alpha^x f(t)e^{m_2 t}\,dt]. \tag{32}$$

Proof. We multiply (31) by $e^{m_1 x}$ and integrate term by term from α to x. This is permissible for an F_2A series, since it is permissible for a Fourier series. We obtain

$$\int_\alpha^x f(t)e^{m_1 t}\,dt = \sum c_k \int_\alpha^x e^{(i\mu_k + m_1)t}\,dt,$$

the series being uniformly convergent. Hence

$$\begin{aligned}\int_\alpha^x f(t)e^{m_1 t}\,dt &= -i[\sum c_k e^{(i\mu_k + m_1)t}/(\mu_k - im_1)]_\alpha^x \\ &= -i \sum c_k e^{(i\mu_k + m_1)x}/(\mu_k - im_1),\end{aligned}$$

in virtue of the first equation (16). Similarly,

$$\int_\alpha^x f(t)e^{m_2 t}\,dt = -i \sum c_k e^{(i\mu_k + m_2)x}/(\mu_k - im_2).$$

Hence (32) equals

$$\begin{aligned}&[2ib/(m_1 - m_2)] \sum c_k e^{i\mu_k x}[m_1/(\mu_k - im_1) - m_2/(\mu_k - im_2)] \\ &\qquad = \sum c_k e^{i\mu_k x} 2ib\mu_k/[(\mu_k - im_1)(\mu_k - im_2)] \\ &= \sum c_k e^{i\mu_k x}(e^{2\pi i\mu_k} - 1), \text{ by (13).}\end{aligned}$$

In the interval $(\alpha+2\pi, \alpha+4\pi)$, (31) is not necessarily an F_2A series. For in order that this shall be the case, it is necessary that the conditions obtained from (16) on replacing α by $\alpha+2\pi$ shall be satisfied. We shall presently see that by altering c_0 and c_1 in various ways, we can ensure that the new series is an F_2A series in $(\alpha+2\pi, \alpha+4\pi)$. It is the F_2A series of its sum $(C,1)$, which is known on account of Theorem 3 and the known changes in c_0 and c_1. Then, by Theorem 3 (applied in the interval $(\alpha+2\pi, \alpha+4\pi)$), we know its sum $(C,1)$ in $(\alpha+4\pi, \alpha+6\pi)$. Hence, we know the sum $(C,1)$ of the original series in $(\alpha+4\pi, \alpha+6\pi)$; and so on.

6. On account of the fact that the coefficients c_k of an F_2A series in $(\alpha, \alpha+2\pi)$ are connected by two equations, it is convenient to omit enough terms to ensure that the constants which remain are independent, and to direct our attention to the remaining series. It would be sufficient to omit terms with suffix $k, -k$ for any assigned $k \neq 0$. The simplest procedure however, seems to be to omit the three central terms with $k = -1, 0, 1$. We are thus led to the following definition.

DEFINITION. *A series*

$$\sum c_k e^{i\mu_k x}, \qquad c_{-k} = \bar{c}_k, \ |k| > 1, \tag{33}$$

is said to be an RF_2A series (reduced F_2A series) in $(\alpha, \alpha+2\pi)$, if there is a function $f(x) \in L(\alpha, \alpha+2\pi)$ such that $c_k = \lambda_k \int_\alpha^{\alpha+2\pi} f(t) e^{-i\mu_k t}\, dt$, $|k| > 1$.

We shall use the notation $\sum''$ to denote a sum in which the values $k = 0, 1, -1$ are omitted. We shall indicate the relation between the RF_2A series in $(\alpha, \alpha+2\pi)$ and the function $f(x)$ by the notation

$$f(x) \approx \sum'' c_k e^{i\mu_k x}, \qquad (\alpha, \alpha+2\pi). \tag{34}$$

THEOREM 4. *If (33) is an RF_2A series in $(\alpha, \alpha+2\pi)$, then it is an RF_2A series in any interval $(\beta, \beta+2\pi)$.*

Proof. It suffices to consider the case in which $\alpha < \beta \leqq \alpha+2\pi$. (See the similar argument in [4], Theorem 1.) Given (34), we define c_r for $r = 0, 1, -1$ by the equation

$$c_r = \lambda_r \int_\alpha^{\alpha+2\pi} f(t) e^{-i\mu_r t}\, dt.$$

Then, by Theorem 3, the series $\sum c_k e^{i\mu_k x}$ is summable $(C,1)$ almost everywhere in $(\alpha+2\pi, \alpha+4\pi)$ to

$$(35)\quad \phi(x) = f(x-2\pi) - 2b/(m_1 - m_2)\cdot Q$$

where $$Q = [m_1 e^{-m_1(x-2\pi)} \int_a^{x-2\pi} f(t)\, e^{m_1 t}\, dt - m_2 e^{-m_2(x-2\pi)} \int_a^{x-2\pi} f(t)\, e^{m_2 t}\, dt].$$

Hence, in the interval $(\beta, \beta+2\pi)$, the series $\sum c_k e^{i\mu_k x}$ is summable (C, 1) almost everywhere to the function $F(x)$ defined by $F(x) = f(x)$ if $\beta \leqq x < \alpha + 2\pi$ and by $F(x) = \phi(x)$ if $\alpha + 2\pi \leqq x < \beta + 2\pi$. We shall now show that there is a complex number d_1 such that, if

$$(36)\quad G(x) = F(x) + d_1 e^{i\mu_1 x} + d_{-1} e^{-i\mu_1 x}, \qquad (d_{-1} = \bar{d}_1),$$

then

$$(37)\quad G(x) \approx \Sigma'' c_k e^{i\mu_k x}, \qquad (\beta, \beta+2\pi).$$

We first evaluate $I = \int_\beta^{\beta+2\pi} F(x)\, e^{-i\mu_k x}\, dx$. We have

$$(38)\quad I = \int_\beta^{\alpha+2\pi} f(x)\, e^{-i\mu_k x}\, dx + \int_{\alpha+2\pi}^{\beta+2\pi} \phi(x)\, e^{-i\mu_k x}\, dx.$$

By (35), the last integral equals

$$e^{-2\pi i\mu_k}[\int_\alpha^\beta f(x) e^{-i\mu_k x}\, dx - (2bm_1/(m_1-m_2)) \int_\alpha^\beta e^{-(m_1+i\mu_k)x}\, dx \int_\alpha^x f(t) e^{m_1 t}\, dt$$
$$+ (2bm_2/(m_1-m_2)) \int_\alpha^\beta e^{-(m_2+i\mu_k)x}\, dx \int_\alpha^x f(t) e^{m_2 t}\, dt].$$

On changing the order of integration in the repeated integrals and using (13), the preceding expression becomes

$$(39)\quad \int_\alpha^\beta f(t) e^{-i\mu_k t}\, dt + J_1(k) - J_2(k),$$

where

$$(40)\quad J_r(k) = [2bm_r/(m_1-m_2)\, e^{-2\pi i\mu_k} e^{-(m_r+i\mu_k)\beta}/(m_r + i\mu_k)] \int_\alpha^\beta f(t)\, e^{m_r t}\, dt.$$

To prove (37), we must show that

$$\int_\beta^{\beta+2\pi} G(x)\, e^{-i\mu_k x}\, dx = \int_\alpha^{\alpha+2\pi} f(x)\, e^{-i\mu_k x}\, dx, \qquad |k| > 1,$$

i. e., that

$$(41)\quad \int_\beta^{\beta+2\pi} (d_1 e^{i\mu_1 x} + d_{-1} e^{-i\mu_1 x})\, e^{-i\mu_k x}\, dx + J_1(k) - J_2(k) = 0, \quad |k| > 1.$$

Now

$$(42)\quad \int_\beta^{\beta+2\pi} e^{i(\mu_1-\mu_k)x}\, dx = e^{i(\mu_1-\mu_k)\beta}/[i(\mu_1-\mu_k)]\, e^{-2\pi i\mu_k}(e^{2\pi i\mu_1} - e^{2\pi i\mu_k}).$$

On subtracting the equation

$$e^{2\pi i\mu_k} - 1 = [2b/(m_1 - m_2)][im_1/(\mu_k - im_1) - im_2/(\mu_k - im_2)]$$

from the similar equation with $k = 1$, and substituting in (42), we find that

$$\int_\beta^{\beta+2\pi} e^{i(\mu_1-\mu_k)x}\,dx = -e^{i(\mu_1-\mu_k)\beta}e^{-2\pi i\mu_k}[2b/(m_1 - m_2)]P,$$

where $P = m_1/(\mu_1 - im_1)(\mu_k - im_1) - m_2/(\mu_1 - im_2)(\mu_k - im_2)$. This equation holds when we replace μ_1 by $-\mu_1$. Substituting these results in (41), the condition to be satisfied takes the form

$$[2b/(m_1 - m_2)]\, e^{-(2\pi+\beta)i\mu_k}[m_1N_1/(\mu_k - im_1) - m_2N_2/(\mu_k - im_2)] = 0, \qquad |k| > 1, \tag{43}$$

where, for $r = 1, 2$,

$$N_r = -d_1e^{i\mu_1\beta}/(\mu_1 - im_r) + d_{-1}e^{-i\mu_1\beta}/(\mu_1 + im_r) - ie^{-m_r\beta}\int_\alpha^\beta f(t)e^{m_rt}\,dt.$$

The condition (43) will therefore be satisfied, if we can choose d_1 so that

$$N_1 = N_2 = 0. \tag{44}$$

If $\int_\alpha^\beta f(t)e^{m_rt}\,dt = 0$, $r = 1, 2$, we can choose $d_1 = 0$. In the contrary case, writing $z_r = d_1e^{i\mu_1\beta}/(\mu_1 - im_r)$, the equations (44) restrict z_r $(r = 1, 2)$ to lie on lines parallel to the real axis. Since $\arg(\mu_1 - im_1) \not\equiv \arg(\mu_1 - im_2)$, $(\mathrm{mod.}\,\pi)$, (44) is equivalent to the condition that d_1 shall lie on two non-parallel lines. This condition can therefore be satisfied.

7. In virtue of Theorem 4, we can use the expression "an RF_2A series" without specifying a particular interval of length 2π.

THEOREM 5. *Given an RF_2A series*

$$\sum'' c_k e^{i\mu_kx} \tag{45}$$

and an interval $(\alpha, \alpha + 2\pi)$, the function $f(x)$ of which (45) is the RF_2A series in $(\alpha, \alpha + 2\pi)$ is determined to within an arbitrary additive multiple of

$$-(a/b)\cdot \sin \pi\mu_1/\mu_1 + \cos\mu_1\,(x - \alpha - \pi).$$

Proof. We may suppose that $\alpha = -\pi$. By definition, there is a function $f_0(x)$ such that

$$c_k = \lambda_k \int_{-\pi}^{\pi} f_0(x)e^{-i\mu_kx}\,dx, \qquad |k| > 1. \tag{46}$$

Since, by Lemma 2, $h(x) = -(a/b) \cdot \sin \pi\mu_1/\mu_1 + \cos \mu_1 x$ is its own F_2A series in $(-\pi, \pi)$, it follows that $c_k = \lambda_k \int_{-\pi}^{\pi} [f_0(x) + th(x)] e^{-i\mu_k x}\, dx$, $|k| > 1$, for all t.

Conversely, let $f(x)$ be a function such that

$$c_k = \lambda_k \int_{-\pi}^{\pi} f(x) e^{-i\mu_k x}\, dx, \qquad (|k| > 1). \tag{47}$$

By (46), $0 = \lambda_k \int_{-\pi}^{\pi} [f(x) - f_0(x)] e^{-i\mu_k x}\, dx$, $(|k| > 1)$. Hence, the F_2A series of $f - f_0$ in $(-\pi, \pi)$ can be written as

$$\gamma_0 + \gamma_1 e^{i\mu_1 x} + \gamma_{-1} e^{-i\mu_1 x}, \qquad (\gamma_{-1} = \bar{\gamma}_1). \tag{48}$$

As in the proof of Theorem 2, γ_1 is real, say $\gamma_1 = \frac{1}{2}t$, and γ_0, γ_1 are connected by (30). Thus, (48) equals $th(x)$. This completes the proof.

Consider now a series $\sum c_k e^{i\mu_k x}$, $(c_{-k} = \bar{c}_k)$, given as the F_2A series of $f(x)$ in $(\alpha, \alpha + 2\pi)$. What changes in c_0, c_1, c_{-1} are necessary and sufficient in order to obtain an F_2A series in $(\alpha + 2\pi, \alpha + 4\pi)$? This problem was referred to at the end of Section 5. By definition, (45) is an RF_2A series in $(\alpha, \alpha + 2\pi)$. By Theorem 4, it is an RF_2A series in $(\alpha + 2\pi, \alpha + 4\pi)$. Hence there are numbers

$$\gamma_0, \gamma_1, \gamma_{-1}, \qquad (\gamma_{-1} = \bar{\gamma}_1), \tag{49}$$

such that

$$\gamma_0 + \gamma_1 e^{i\mu_1 x} + \gamma_{-1} e^{-i\mu_1 x} + \sum'' c_k e^{i\mu_k x} \tag{50}$$

is an F_2A series in $(\alpha + 2\pi, \alpha + 4\pi)$. By Theorem 5, with α replaced by $\alpha + 2\pi$, there is a unique triad of numbers (49) with $\gamma_0 = 0$; say $0, \gamma_1^*, \gamma_{-1}^*$. Write, for convenience, $\gamma_k^* = c_k$, $(|k| > 1)$. Then γ_1^* is uniquely determined by the appropriate form of equations (16); viz.

$$\sum \gamma_k^* e^{i\mu_k(\alpha+2\pi)}/(\mu_k - im_1) = 0, \qquad \sum \gamma_k^* e^{i\mu_k(\alpha+2\pi)}/(\mu_k - im_2) = 0.$$

It follows from Theorem 5, with α replaced by $\alpha + 2\pi$, that the formulae for $\gamma_0, \gamma_1, \gamma_{-1}$ are

$$-t(a/b) \cdot \sin \pi\mu_1/\mu_1, \qquad \gamma_1^* + \tfrac{1}{2}t\, e^{-i\mu_1(\alpha+3\pi)}, \qquad \gamma_{-1}^* + \tfrac{1}{2}t\, e^{i\mu_1(\alpha+3\pi)}.$$

8. Another simple consequence of Theorem 5 is the following lemma, which will be used in the proof of the Young-Hausdorff inequalities for F_2A series.

LEMMA 3. *If L, N, α are real numbers, and M is a complex number such that*

$$-L(a/b)\cdot \sin \pi\mu_1/\mu_1 + \operatorname{Re} Me^{-i\mu_1(\alpha+\pi)} \neq 0, \tag{51}$$

then, given an RF_2A series (45), there is precisely one triad c_0, c_1, c_{-1} with c_0 real, $c_{-1} = \bar{c}_1$, such that

$$Lc_0 + Mc_1 + \bar{M}c_{-1} = N, \tag{52}$$

and such that $\sum c_k e^{i\mu_k x}$ is an F_2A series in $(\alpha, \alpha + 2\pi)$.

Proof. There are numbers (49) such that (50) is an F_2A series in $(\alpha, \alpha + 2\pi)$. By Theorem 5, the general formulae for c_0, c_1, c_{-1}, such that $\sum c_k e^{i\mu_k x}$ is an F_2A series in $(\alpha, \alpha + 2\pi)$, are

$$c_0 = \gamma_0 - t(a/b)\cdot \sin \pi\mu_1/\mu_1, \quad c_1 = \gamma_1 + \tfrac{1}{2}t\, e^{-i\mu_1(\alpha+\pi)}, \quad c_{-1} = \gamma_{-1} + \tfrac{1}{2}t\, e^{i\mu_1(\alpha+\pi)}.$$

On substituting these values in (52), we obtain a linear equation in t. It has a unique solution, since the coefficient of t is the first member of (51).

9. The results of this section are proved in the same way as the analogous results concerning F. A. series.

THEOREM 6. *If $f(x) \sim \sum c_k e^{i\mu_k x}$, $(\alpha, \alpha + 2\pi)$, and $f \in L^2(\alpha, \alpha + 2\pi)$, then $\int_\alpha^{\alpha+2\pi} f^2(x)\,dx = \sum |c_\nu|^2/\lambda_\nu$.*

COROLLARY. *If F is a complex function, and $F \in L^2(\alpha, \alpha + 2\pi)$, and*

$$d_\nu = \lambda_\nu \int_\alpha^{\alpha+2\pi} F(t) e^{-i\mu_\nu t}\,dt, \text{ then } \sum |d_\nu|^2/\lambda_\nu = \int_\alpha^{\alpha+2\pi} |F(t)|^2\,dt.$$

Proof. As for [4], Theorem 2, and Corollary.

LEMMA 4. *If (45) is an RF_2A series, then it is summable (C, 1) almost everywhere to a function $g(x)$; and, if $v(x)$ is a function of bounded variation in an assigned interval (a, b), then*

$$\lim_{n\to\infty} \int_a^x v(t) \sum_{|\nu|<n}'' c_\nu e^{i\mu_\nu t}\,dt = \int_a^x v(t)g(t)\,dt,$$

as $n \to \infty$, uniformly for $a \leqq x \leqq b$.

Proof. As for [4], Lemma 6.

THEOREM 7. *If $g(x)$ is the sum (C, 1) of the RF_2A series (45), then the series is the Fourier series of $g(x)$; i.e., if $\mu \geqq 0$, then*

$$\lim_{|T|\to\infty} 1/T \int_0^T g(x)e^{-i\mu x}dx = c_s \text{ if } \mu = \mu_s, s = 2, 3, \cdots,$$
$$= 0 \text{ otherwise.}$$

Proof. We need to know that if $\mu, \geqq 0$, is assigned, then

$$\sum_{|k|>s} c_k/(\mu_k - \mu) \tag{53}$$

is convergent, s being a positive integer greater than 1. By Theorems 1 and 4, $\sum'' c_k e^{i\mu_k \alpha}/(\mu_k - im_1)$ is convergent for all α. In particular,

$$\sum_{|k|>s} c_k/(\mu_k - im_1) \tag{54}$$

is convergent. On subtracting (54) from (53), and using $c_k = o(1)$, we see that (53) is convergent. The proof proceeds as in [4], Theorem III. Note that here, $\mu_k = k - 1 + \eta_k$, while in [4], $\mu_k = k - \frac{1}{2} + \eta_k$, with $k\eta_k = O(1)$.

LEMMA 5. *If $d_0, d_1, d_{-1}, \cdots, d_m, d_{-m}$ are any complex numbers and $1 < p \leqq 2$, $p^{-1} + q^{-1} = 1$, then*

$$\left[\int_\alpha^{\alpha+2\pi} \left|\sum_{|\nu|\leqq m} \lambda_\nu^{1/q} d_\nu e^{i\mu_\nu x}\right|^q dx\right]^{1/q} \leqq \left[\sum_{|\nu|\leqq m} |d_\nu|^p\right]^{1/p}.$$

Proof. As for [4], Lemma 9.

10. We now consider the Young-Hausdorff inequalities for F_2A series. On account of the fact that the F_2A constants are connected by two equations, we cannot assign c_ν for all integer ν subject to $c_{-\nu} = \bar{c}_\nu$, $\sum |c_\nu|^p < \infty$, $1 < p \leqq 2$, and expect the c_ν to be the F_2A constants of a real function of class $L^q(\alpha, \alpha + 2\pi)$. On the other hand, if we omit c_0, c_1, c_{-1}, then there is one degree of freedom in the set of possible functions of which the c_ν are the F_2A constants. If, then, we require that the unspecified values of the three central constants shall satisfy an assigned linear equation, we may expect to find a unique function satisfying all the conditions.

THEOREM 8. (a) *Let c_ν, $|\nu| > 1$ be complex numbers such that $\Sigma''|c_\nu|^p < \infty$ and $c_{-\nu} = \bar{c}_\nu$. Let L, N, α be assigned real numbers, and let M be an assigned complex number such that* (51) *holds. Then there is precisely one (real) function $f \in L^q(\alpha, \alpha + 2\pi)$ which satisfies the following two conditions:*

(i) $$c_\nu = \lambda_\nu \int_\alpha^{\alpha+2\pi} f(t)e^{-i\mu_\nu t}dt, \qquad |\nu| > 1;$$

(ii) *If c_0, c_1, c_{-1} are defined by the preceding equation for $\nu = 0, 1, -1$, then $Lc_0 + Mc_1 + \bar{M}c_{-1} = N$. For this function,*

$$(55) \qquad \left(\int_\alpha^{\alpha+2\pi} |f|^q dx\right)^{1/q} \leqq \left(\sum |c_\nu|^p/\lambda_\nu^{p-1}\right)^{1/p}.$$

(b) *If $g(x)$ is a real function of class $L^p(\alpha, \alpha+2\pi)$, and*

$$\gamma_\nu = \lambda_\nu \int_\alpha^{\alpha+2\pi} g(x) e^{-i\mu_\nu x} dx, \qquad (\nu = 0, \pm 1, \cdots),$$

then

$$\left(\sum |\gamma_\nu|^q/\lambda_\nu^{q-1}\right)^{1/q} \leqq \left(\int_\alpha^{\alpha+2\pi} |g|^p dx\right)^{1/p}.$$

Proof of (a). We define $c_0(n)$, $c_1(n)$ and $c_{-1}(n)$ by the equations

$$(56) \qquad Lc_0(n) + Mc_1(n) + \bar{M}c_{-1}(n) = N,$$

$$(57) \qquad c_0(n)/(-im_1) + c_1(n)e^{i\mu_1\alpha}/(\mu_1 - im_1) + c_{-1}(n)e^{-i\mu_1\alpha}/(-\mu_1 - im_1) = -\sum_{|k|\leqq n}'' c_k e^{i\mu_k\alpha}/(\mu_k - im_1),$$

$$(58) \qquad c_0(n)/(-im_2) + c_1(n)e^{i\mu_1\alpha}/(\mu_1 - im_2) + c_{-1}(n)e^{-i\mu_1\alpha}(-\mu_1 - im_2) = -\sum_{|k|\leqq n}'' c_k e^{i\mu_k\alpha}/(\mu_k - im_2).$$

The determinant of this system of equations is

$$(i \sin \pi\mu_1/ab)(m_1 - m_2)\{-L(2a/b)\cdot \sin \pi\mu_1/\mu_1 + Me^{-i\mu_1(\alpha+\pi)} + \bar{M}e^{i\mu_1(\alpha+\pi)}\} \neq 0.$$

The second members of (57), (58) are purely imaginary (or zero). Multiply (57) by im_1, (58) by im_2, and subtract. We obtain an equation of the form $c_1(n)A + c_{-1}(n)\bar{A} = B$, where B is real and $A \neq 0$. Hence $c_{-1}(n) = \bar{c}_1(n)$. Hence, in (57), all terms except the first are purely imaginary. Hence $c_0(n)$ is real. Since the series

$$\sum'' c_k e^{i\mu_k\alpha}/(\mu_k - im_1), \qquad \sum'' c_k e^{i\mu_k\alpha}(\mu_k - im_2),$$

converge with $\sum''|c_\nu|^p$, it follows that there are numbers γ_0 (real), γ_1, γ_{-1} $(= \bar{\gamma}_1)$, such that $c_0(n) \to \gamma_0$, $c_1(n) \to \gamma_1$, $c_{-1}(n) \to \gamma_{-1}$, as $n \to \infty$. By (56),

$$(59) \qquad L\gamma_0 + M\gamma_1 + \bar{M}\gamma_{-1} = N.$$

Let $T_n(x) = c_0(n) + c_1(n)e^{i\mu_1 x} + c_{-1}(n)e^{-i\mu_1 x} + \sum_{|k|\leqq n}'' c_k e^{i\mu_k x}$. By Lemma 5, $\{T_n(x)\}$ converges strongly with index q to a real function, f say. By Theorem 2, and (57), (58), $T_n(x)$ is its own F_2A series in $(\alpha, \alpha+2\pi)$. Hence, for a given integer ν with $|\nu| > 1$ and for $n > |\nu|$,

$$c_\nu/\lambda_\nu = \int_\alpha^{\alpha+2\pi} T_n(x)\, e^{-i\mu_\nu x} dx = \lim_{n\to\infty} \int_\alpha^{\alpha+2\pi} T_n(x) e^{-i\mu_\nu x} dx$$

$$= \int_\alpha^{\alpha+2\pi} f(x) e^{-i\mu_\nu x} dx.$$

Further, for $\nu = 0, 1, -1$,

$$\gamma_\nu = \lim_{n\to\infty} c_\nu(n) = \lim_{n\to\infty} \lambda_\nu \int_\alpha^{\alpha+2\pi} T_n(x) e^{-i\mu_\nu x} dx = \lambda_\nu \int_\alpha^{\alpha+2\pi} f(x) e^{-i\mu_\nu x} dx = c_\nu.$$

Thus, the conditions (i), (ii) are satisfied. Further, since, by Lemma 5,

$$\left(\int_\alpha^{\alpha+2\pi} |T_n(x)|^q dx\right)^{1/q} \leqq [|c_0(n)|^p/\lambda_0^{p-1} + |c_1(n)|^p/\lambda_1^{p-1}$$

$$+ |c_{-1}(n)|^p/\lambda_1^{p-1} + \sum''_{|\nu|\leqq n} |c_\nu|^p/\lambda_\nu^{p-1}]^{1/p},$$

(55) follows on letting $n \to \infty$.

That there is only one function $f(x)$ which satisfies the conditions (i), (ii), follows from Lemma 3. For if f_0 is such a function, then, by Lemma 3, f_0 has the same F_2A constants as f. The F_2A series of $f_0 - f$ is thus identically zero. But $f_0 - f$ is, almost everywhere in $(\alpha, \alpha + 2\pi)$, the sum (C, 1) of its F_2A series. Hence $f_0 - f = 0$.

Proof of (b). As for [4], Theorem IV(b).

11. We now consider the simplest theorem of uniqueness for convergent series of the form $\sum c_k e^{i\mu_k x}$ with $c_{-k} = \bar{c}_k$. It is no loss of generality to suppose that the interval of convergence, given as of length 2π, is $(-\pi, \pi)$.

Consider the F_2A series of the function $f(x)$ defined as equal to 1 for $|x| \leqq \pi$. Then $1 \sim \sum \gamma_k e^{i\mu_k x}$, where $\gamma_0 = 2\pi\lambda_0$, $\gamma_k = 2\lambda_k \sin \pi\mu_k/\mu_k$, $(k \neq 0)$. By (1), $A_k^2/\cos^2 \pi\mu_k = B_k^2/\sin^2 \pi\mu_k = A_k^2 + B_k^2$. Hence

$$\mu_k^2 \sin^2 \pi\mu_k = b^2\mu_k^4/[(\mu_k^2 + a)^2 + b^2\mu_k^2] \to b^2, \tag{60}$$

as $|k| \to \infty$. It follows that

$$\gamma_0 - 1 + \sum' \gamma_k e^{i\mu_k x} \tag{61}$$

is a series whose coefficients are $O(k^{-2})$, and which converges to zero in the closed interval $(-\pi, \pi)$. We thus see that in order to obtain a significant uniqueness theorem, we must introduce some explicit restriction on the coefficients. It will be seen that the condition $c_0 = 0$ is such a restriction.

LEMMA 6. *If* $\sum' c_k e^{i\mu_k x}$, *where* $c_{-k} = \bar{c}_k$, $kc_k = o(1)$, *converges to zero almost everywhere in* $(-\pi, \pi)$, *then* $c_k = 0$ $(k = \pm 1, \pm 2, \cdots)$.

Proof. By hypothesis, $\Sigma'' c_k e^{i\mu_k x}$ converges almost everywhere in $(-\pi, \pi)$ to $-c_1 e^{i\mu_1 x} - c_{-1} e^{-i\mu_1 x}$. In Theorem 8, let $p=2$, $M=N=0$, $\alpha=-\pi$, $L=1$. By the first part of that theorem, there are numbers c_1^*, c_{-1}^* $(=\bar{c}_1^*)$ such that

$$c_1^* e^{i\mu_1 x} + c_{-1}^* e^{-i\mu_1 x} + \Sigma'' c_k e^{i\mu_k x} \tag{62}$$

is an F_2A series in $(-\pi, \pi)$. But this series converges almost everywhere in $(-\pi, \pi)$ to

$$(c_1^* - c_1) e^{i\mu_1 x} + (c_{-1}^* - c_{-1}) e^{-i\mu_1 x}. \tag{63}$$

Hence (62) is the F_2A series of (63) in $(-\pi, \pi)$. Write

$$c_1^* - c_1 = \tfrac{1}{2}(\alpha - i\beta). \tag{64}$$

Then (63) is $\alpha \cos \mu_1 x + \beta \sin \mu_1 x$. Since the constant term in (62) is zero,

$$0 = \lambda_0 \int_{-\pi}^{\pi} (\alpha \cos \mu_1 x + \beta \sin \mu_1 x)\, dx.$$

Hence, $\alpha = 0$. For $|k| > 1$, we have

$$c_k/\lambda_k = \int_{-\pi}^{\pi} \beta \sin \mu_1 x\, e^{-i\mu_k x}\, dx = 2i\beta/b \cdot \sin \pi\mu_1 \sin \pi\mu_k,$$

by (20). By (7) and (60), $|\mu_k c_k| \to \frac{1}{\pi} |\beta \sin \pi\mu_1|$. By hypothesis, $\mu_k c_k \to 0$. Further, by (6), μ_1 is not an integer. Hence $\beta = 0$. Thus, all the coefficients in (62) are zero. By (64), $c_1 = c_1^*$. Hence all the c_k are zero.

The series obtained by differentiating (61) term by term, viz.

$$\Sigma' i\mu_k \gamma_k e^{i\mu_k x}, \tag{65}$$

has its coefficients $O(k^{-1})$. As in the proof of Lemma 6, if we alter the two central terms suitably, we obtain an F_2A series, say

$$\phi(x) \sim i\mu_1 \gamma_1^* e^{i\mu_1 x} - i\mu_1 \gamma_{-1}^* e^{-i\mu_1 x} + \Sigma'' i\mu_k \gamma_k e^{i\mu_k x}, \qquad (-\pi, \pi). \tag{66}$$

This series can be integrated term by term from 0 to x. Since (61) converges to zero, it follows that

$$\int_0^x \phi(u)\, du - (\gamma_1^* - \gamma_1) e^{i\mu_1 x} - (\gamma_{-1}^* - \gamma_{-1}) e^{-i\mu_1 x}$$

is constant in $|x| \leqq \pi$. Hence we may take $\phi(x)$ to be

$$i\mu_1 (\gamma_1^* - \gamma_1) e^{i\mu_1 x} - i\mu_1 (\gamma_{-1}^* - \gamma_{-1}) e^{-i\mu_1 x},$$

for $|x| \leqq \pi$. On account of the equiconvergence of an F_2A series in $(-\pi, \pi)$ with a Fourier series in $(-\pi, \pi)$, the second member of (66) converges to

this function, and so (65) converges to zero for $|x| \leqq \pi$. That it is essentially the only series without a constant term which has this property, is shown by the following theorem.

THEOREM 9. *If $\sum' c_k e^{i\mu_k x}$, where $c_{-k} = \bar{c}_k$, converges to zero for $|x| < \pi$, then there is a real constant c such that $c_k = ci\mu_k\gamma_k$.*

Proof. The hypothesis of convergence implies that $c_k = o(1)$. Let

$$F(x) = -\sum c_k/\mu_k^2 \cdot e^{i\mu_k x}. \tag{67}$$

By the classical argument of Riemann, $D^2F(x) = 0$ for $|x| < \pi$, where $D^2F(x)$ denotes the generalised second derivative. Hence, $F(x) = lx + m$, where l, m are real numbers.

With γ_k defined as above, write $d_k = \gamma_k/(1-\gamma_0)$, $(k \neq 0)$. Then

$$\sum' d_k e^{i\mu_k x} = 1, \qquad (|x| \leqq \pi), \tag{68}$$

and $|\mu_k^2 d_k| = |2\lambda_k\mu_k \sin \pi\mu_k/(1-\gamma_0)| \to |b/\pi(1-\gamma_0)|$, as $|k| \to \infty$. The series (68) is uniformly convergent. Hence, $\sum' d_k/i\mu_k \cdot e^{i\mu_k x} = x$ $(|x| \leqq \pi)$, and

$$\sum' d_k(l/(i\mu_k) + m)e^{i\mu_k x} = lx + m. \tag{69}$$

The difference between the series in (69) and (67) converges to zero for $|x| < \pi$, and has coefficients which are $O(k^{-2})$. By Lemma 6, they are all zero. Thus $-c_k/\mu_k^2 = d_k(m + l/i\mu_k)$. Multiply by μ_k^2 and let $k \to \infty$. Then $0 = |mb/[\pi(1-\gamma_0)]|$. Thus, $m = 0$, and $c_k = ild_k\mu_k = l/(1-\gamma_0) \cdot i\mu_k\gamma_k$, as required.

Sections 12-16. F_2B Series.

12. The theory of F_2B series can be developed on lines similar to the preceding theory of F_2A series. The details of the analysis, however, are sufficiently different to justify the short sketch which follows. The notation will be almost the same as for F_2A series, but the meanings of the symbols will be different.

The F_2B series (see Section 1) is a series of cosines and sines of $\mu_k x$ where the μ_k are the roots, supposed all real and simple, of the equation

$$A \sin \pi z - B \cos \pi z = 0, \tag{70}$$

where $A = az$, $B = z^2 + b$, $ab \neq 0$. Then $z = 0$ is not a root of (70).

If $f(x)$ is a real function of class $L(\alpha, \alpha + 2\pi)$, then the F_2B series of $f(x)$ in $(\alpha, \alpha + 2\pi)$ can be written as

$$f(x) \sim \sum' c_k e^{i\mu_k x}, \qquad (\alpha, \alpha + 2\pi), \tag{71}$$

where $c_k = \lambda_k \int_\alpha^{\alpha+2\pi} f(x) e^{-i\mu_k x} dx$, $(k = \pm 1, \pm 2, \cdots)$, and $\mu_1 < \mu_2 < \cdots$ are the positive roots of the equation (70), $\mu_{-k} = -\mu_k$, $c_{-k} = \bar{c}_k$, and λ_k is the residue at $z = \mu_k$ of

$$Q(z) = \tfrac{1}{2}(A \cos \pi z + B \sin \pi z)/(A \sin \pi z - B \cos \pi z).$$

We find that

$$\mu_k = k - \tfrac{3}{2} + \eta_k, \qquad k\eta_k = O(1). \tag{72}$$

Writing $A_k = A(\mu_k)$, etc., we have

$$1/(2\lambda_k) = \pi + (A_k' \sin \pi\mu_k - B_k' \cos \pi\mu_k)/(A_k \cos \pi\mu_k + B_k \sin \pi\mu_k).$$

But $A_k/\cos \pi\mu_k = B_k/\sin \pi\mu_k = A_k \cos \pi\mu_k + B_k \sin \pi\mu_k$. Hence

$$1/(2\lambda_k) = \pi + (a \sin \pi\mu_k - 2\mu_k \cos \pi\mu_k)\cos \pi\mu_k/a\mu_k. \tag{73}$$

If ν, n are positive integers, we have

$$\int_{-\pi}^{\pi} \cos \mu_\nu x \cos \mu_n x \, dx = (2/a) \cdot \cos \pi\mu_\nu \cos \pi\mu_n, \qquad (n \neq \nu), \tag{74}$$

$$\int_{-\pi}^{\pi} \sin \mu_\nu x \sin \mu_n x \, dx = (-2b/a) \cdot (\cos \pi\mu_\nu/\mu_\nu) \cdot (\cos \pi\mu_n/\mu_n), \quad (n \neq \nu). \tag{75}$$

13. The F_2B series (71) is uniformly equiconvergent in the closed interval $(\alpha, \alpha + 2\pi)$ with a Fourier series, though not with the Fourier series of a function of period 2π. We have the following special case of a known theorem ([3], Theorem 1).

LEMMA 7. *If $s_n(x)$ is the n-th partial sum of* (71), *i. e.*, $s_n = \sum'_{|k| \leqq n} c_k e^{i\mu_k x}$, *and if $\tau_{2n}(x)$ is the 2n-th partial sum of the Fourier series of the function $\phi(x)$, of period 4π, defined by $\phi(x) = f(x)$ if $\alpha \leqq x \leqq \alpha + 2\pi$ and by $\phi(x) = -f(x - 2\pi)$ if $\alpha + 2\pi < x < \alpha + 4\pi$, then $s_n(x) - \tau_{2n}(x) \to 0$ uniformly in $(\alpha, \alpha + 2\pi)$ as $n \to \infty$. In particular, $s_n(\alpha + 2\pi) + s_n(\alpha) \to 0$ as $n \to \infty$.*

According to the last clause of Lemma 7,

$$\sum' c_k e^{i\mu_k x}(e^{2\pi i\mu_k} + 1) = 0. \tag{76}$$

Equation (70) can be written $e^{2\pi iz} = (A + iB)/(A - iB)$. Hence

$$e^{2\pi i\mu_k} + 1 = 2ia\mu_k/(\mu_k^2 + ia\mu_k + b) = 2ia\mu_k/[(\mu_k + im_1)(\mu_k + im_2)]. \tag{77}$$

We shall suppose that $a^2 + 4b > 0$, so that m_1, m_2 are real and different. We shall now prove the analogue of Theorem 1.

THEOREM 10. *For any F_2B series* (71), *we have*

$$\sum' c_k e^{i\mu_k \alpha} \mu_k / [(\mu_k + im_1)(\mu_k + im_2)] = 0, \tag{78}$$

and

$$\sum' c_k e^{i\mu_k \alpha} / [(\mu_k + im_1)(\mu_k + im_2)] = 0, \tag{79}$$

or, what is equivalent,

$$\sum' c_k e^{i\mu_k \alpha} / (\mu_k + im_1) = 0, \qquad \sum' c_k e^{i\mu_k \alpha} / (\mu_k + im_2) = 0. \tag{80}$$

Proof. Equation (78) follows from (76) on substituting the value of $e^{2\pi i \mu_k} + 1$ from (77). To prove (79), or, what is equivalent,

$$\sum' c_k / \mu_k \cdot e^{i\mu_k \alpha} (e^{2\pi i \mu_k} + 1) = 0, \tag{81}$$

it suffices to show that

$$\sum' \lambda_k / \mu_k \cdot \cos \pi \mu_k e^{i\mu_k (x - \alpha - \pi)} = 0, \tag{82}$$

uniformly in $(\alpha, \alpha + 2\pi)$. For then, on taking conjugates, multiplying by $f(x)$ and integrating over $(\alpha, \alpha + 2\pi)$, we obtain (81).

Changing the variable to $x - \alpha - \pi$, we have to prove that

$$\sum_1^\infty \lambda_n / \mu_n \cdot \cos \pi \mu_n \sin \mu_n x = 0, \tag{83}$$

uniformly in $(-\pi, \pi)$.

Consider the F_2B series, in the interval $(-\pi, \pi)$, of the function $\sin \mu_1 x$; say

$$\sum_1^\infty b_k \sin \mu_k x. \tag{84}$$

For $k > 1$, we have by (75),

$$b_k = -4b\lambda_k / a \cdot \cos \pi \mu_1 / \mu_1 \cdot \cos \pi \mu_k / \mu_k. \tag{85}$$

Further,

$$\begin{aligned} b_1 &= 2\lambda_1 (\pi - 1/\mu_1 \cdot \sin \pi \mu_1 \cos \pi \mu_1), \\ b_1 - 1 &= 2\lambda_1 (\pi - 1/(2\lambda_1) - 1/\mu_1 \cdot \sin \pi \mu_1 \cos \pi \mu_1) \\ &= 2\lambda_1 (-2 \cos \pi \mu_1 / \mu_1)(\sin \pi \mu_1 - \mu_1 / a \cdot \cos \pi \mu_1) \\ &= -4b\lambda_1 / a (\cos \pi \mu_1 / \mu_1)^2, \end{aligned} \tag{86}$$

by (73) and (70). By Lemma 7, the series (84) converges uniformly to $\sin \mu_1 x$. Substituting the values (85), (86), we obtain (83). Clearly, (78), (79), are equivalent to (80).

14. LEMMA 8. *If ν is a positive integer > 1, then*

$$f_\nu(x) = -\cos \mu_1 x/\cos \mu_1 \pi + \cos \mu_\nu x/\cos \mu_\nu \pi,$$

is its own F_2B series in $(-\pi, \pi)$; and

$$g_\nu(x) = -\mu_1 \sin \mu_1 x/\cos \mu_1 \pi + \mu_\nu \sin \mu_\nu x/\cos \mu_\nu \pi,$$

is its own F_2B series in $(-\pi, \pi)$.

Proof. The F_2B series of $f_\nu(x)$ can be written $\sum_1^\infty a_n \cos \mu_n x$. If $n \neq 1, \nu$, then $a_n = 2\lambda_n \int_{-\pi}^{\pi} f_\nu(x) \cos \mu_n x \, dx = 0$, by (74). Hence the F_2B series of $f_\nu(x)$ is $a_1 \cos \mu_1 x + a_\nu \cos \mu_\nu x$. This converges to $f_\nu(x)$ for $|x| < \pi$. Giving x the value $\pi/2\mu_\nu$, we find $a_1 = -\sec \mu_1 \pi$. Hence $a_\nu = \sec \mu_\nu \pi$. The proof of the second part is similar.

THEOREM 11. *In order that*

$$\sum_{|k| \leqq n}{}' c_k e^{i\mu_k x}, \qquad c_{-k} = \bar{c}_k, \tag{87}$$

shall be an F_2B series in $(\alpha, \alpha + 2\pi)$, it is necessary and sufficient that

$$\sum_{|k| \leqq n}{}' c_k e^{i\mu_k \alpha}(e^{2\pi i \mu_k} + 1) = 0, \tag{88}$$

and that

$$\sum_{|k| \leqq n}{}' c_k/\mu_k \cdot e^{i\mu_k \alpha}(e^{2\pi i \mu_k} + 1) = 0. \tag{89}$$

Proof. By Theorem 10 and (77), the conditions (88), (89) are necessary. To prove that they are sufficient, we may suppose that $\alpha = -\pi$. Then the conditions become, on using the notation $c_k = \frac{1}{2}(a_k - ib_k)$,

$$\sum_1^n a_k \cos \pi\mu_k = 0, \qquad \sum_1^n b_k/\mu_k \cdot \cos \pi\mu_k = 0.$$

Hence $\sum_1^n a_k \cos \mu_k x + b_k \sin \mu_k x$ becomes

$$\sum_2^n a_k(\cos \mu_k x - \cos \mu_k \pi/(\cos \mu_1 \pi) \cdot \cos \mu_1 x) + \sum_2^n b_k(\sin \mu_k x - \mu_1 \cos \mu_k \pi/(\mu_k \cos \mu_1 \pi) \cdot \sin \mu_1 x).$$

By Lemma 8, each term is an F_2B series in $(-\pi, \pi)$. So then is the first member, i.e. (87).

Enough has now been said to make it clear that F_2B series may be investigated in the same way as F_2A series. Some results are simpler, since,

in an F_2B series, c_1, c_{-1} are determined by the remaining coefficients, and there is no constant term. We shall not enunciate the analogues of Theorems 3-8. We shall merely enunciate a special case of the analogue of Theorem 8(a), which will be used below.

LEMMA 9. *Let c_ν, $|\nu|>1$, be complex numbers such that $\sum''|c_\nu|^2<\infty$, $c_{-\nu}=\bar{c}_\nu$. Given the interval $(\alpha,\alpha+2\pi)$, there is precisely one real function f of class $L^2(\alpha,\alpha+2\pi)$, such that $c_\nu=\lambda_\nu\int_\alpha^{\alpha+2\pi}f(t)e^{-i\mu_\nu t}dt$, $(|\nu|>1)$.*

15. We now turn to the problem of the uniqueness of convergent series of the form $\sum' c_k e^{i\mu_k x}$ with $c_{-k}=\bar{c}_k$. We shall prove (Theorem 12) that any series of this form which converges to zero in the open interval $(-\pi,\pi)$, is a linear combination of two particular series with this property. The particular series are (83), and the series obtained from it by termwise differentiation. We shall use the fact that, since $\tan\pi\mu_k=(\mu_k^2+b)/a\mu_k$, we have, by (72),

$$\tan\pi\eta_k=-a\mu_k/(\mu_k^2+b)=-a/\mu_k+O(k^{-3}),$$

$$\eta_k=-a/\pi\mu_k+O(k^{-3}),$$

$$\cos\pi\mu_k=(-1)^{k-1}\sin\pi\eta_k=(-1)^k a/\mu_k+O(k^{-3}),$$

and

$$|\mu_k\cos\pi\mu_k|\to|a|, \qquad \text{as } k\to\infty. \tag{90}$$

16. LEMMA 10. *If*

$$\sum' c_k e^{i\mu_k x}, \qquad c_{-k}=\bar{c}_k,\ kc_k=o(1), \tag{91}$$

converges uniformly to zero in $(-\pi,\pi)$, then there is a real number l such that

$$c_k=l(i\lambda_k/\mu_k)\cdot\cos\pi\mu_k; \tag{92}$$

i.e., (91) *is a constant multiple of* (83).

Proof. The series $\sum'' c_k e^{i\mu_k x}$ converges uniformly to $-c_1e^{i\mu_1 x}-c_{-1}e^{-i\mu_1 x}$. By Lemma 9, there are numbers c_1^*, c_{-1}^* $(=\bar{c}_1^*)$ such that

$$c_1^*e^{i\mu_1 x}+c_{-1}^*e^{-i\mu_1 x}+\sum'' c_k e^{i\mu_k x}$$

is an F_2B series in $(-\pi,\pi)$. It is the F_2B series of its sum, viz.,

$$(c_1^*-c_1)e^{i\mu_1 x}+(c_{-1}^*-c_{-1})e^{-i\mu_1 x}.$$

Write $c_1^* - c_1 = \frac{1}{2}(\alpha - i\beta)$. Then, for $|k| > 1$,

$$(93) \qquad c_k/\lambda_k = \int_{-\pi}^{\pi} (\alpha \cos \mu_1 x + \beta \sin \mu_1 x) e^{-i\mu_k x}\, dx$$

$$= \alpha \int_{-\pi}^{\pi} \cos \mu_1 x \cos \mu_k x\, dx - i\beta \int_{-\pi}^{\pi} \sin \mu_1 x \sin \mu_k x\, dx$$

$$= 2\alpha/a \cdot \cos \mu_1 \pi \cos \mu_k \pi + 2ib\beta/a \cdot \cos \mu_1 \pi/\mu_1 \cdot \cos \mu_k \pi/\mu_k,$$

by (74), (75). By hypothesis, $kc_k \to 0$. By (70), $\cos \pi\mu_1 \neq 0$. Multiply (93) by μ_k, and let $k \to \infty$. By (90), $\alpha = 0$. Hence we have (92) for $|k| > 1$, with $l = 2b\beta/a \cdot \cos \mu_1 \pi/\mu_1$. Since $c_1 = \frac{1}{2}(a_1 - ib_1)$, (91) is

$$a_1 \cos \mu_1 x + b_1 \sin \mu_1 x - 2l \sum_2^{\infty} \lambda_n/\mu_n \cdot \cos \mu_n \pi \sin \mu_n x,$$

and converges to $a_1 \cos \mu_1 x + b_1 \sin \mu_1 x + 2l(\lambda_1/\mu_1) \cdot \cos \mu_1 \pi \sin \mu_1 x$, by (83). By hypothesis, this is zero. Hence $a_1 = 0$, $b_1 = -2l(\lambda_1/\mu_1) \cdot \cos \mu_1 \pi$, as required.

LEMMA 11. *The series*

$$(94) \qquad \sum' \lambda_k \cos \pi\mu_k\, e^{i\mu_k x}$$

converges boundedly to zero in the open interval $(-\pi, \pi)$.

Proof. By Lemma 9, there are numbers λ_1^*, λ_{-1}^* $(= \bar{\lambda}_1^*)$ such that

$$(95) \qquad \lambda_1^* e^{i\mu_1 x} + \lambda_{-1}^* e^{-i\mu_1 x} + \sum'' \lambda_k \cos \pi\mu_k e^{i\mu_k x}$$

is an F_2B series in $(-\pi, \pi)$. Hence it is summable (C, 1) almost everywhere in that interval, and can be integrated term by term. The same is therefore true of (94). Denote the sum (C, 1) of (94) by $f(x)$. Then

$$\sum' \lambda_k/(i\mu_k) \cdot \cos \pi\mu_k e^{i\mu_k x} = \int_0^x f(t)\, dt.$$

By (83) the first member is zero. Hence $f = 0$ almost everywhere. Then (95) is the F_2B series of

$$(96) \qquad (\lambda_1^* - \lambda_1) e^{i\mu_1 x} + (\lambda_{-1}^* - \lambda_{-1}) e^{-i\mu_1 x}.$$

By Lemma 7, (95) converges boundedly in the open interval $(-\pi, \pi)$ to (96). Hence (94) converges boundedly in this interval to zero.

We finally have,

THEOREM 12. *If* $\sum' c_k e^{i\mu_k x}$, *where* $c_{-k} = \bar{c}_k$, *converges to zero in the open interval* $(-\pi, \pi)$, *then it is a linear combination of the two series*

$$\sum' \lambda_k \cos \pi\mu_k e^{i\mu_k x}, \qquad \sum' \lambda_k/(i\mu_k) \cdot \cos \pi\mu_k e^{i\mu_k x},$$

of which the first converges boundedly, and the second converges uniformly to zero in the open interval.

Proof. As for Theorem 9, $-\sum' c_k/\mu_k^2 \cdot e^{i\mu_k x} = lx + m$ with real constant l and m. On integrating (83), we infer that $\sum' \lambda_k/\mu_k^2 \cdot \cos \pi\mu_k e^{i\mu_k x}$ converges uniformly to a constant p. By Lemma 10, this constant is not zero. Then $\sum' \lambda_k/(i\mu_k^3) \cdot \cos \pi\mu_k e^{i\mu_k x}$ converges uniformly to px. Hence

$$\sum [l/p \cdot \lambda_k/(i\mu_k^3) \cdot \cos \pi\mu_k + m/p \cdot \lambda_k/\mu_k^2 \cos \pi\mu_k + c_k/\mu_k^2] e^{i\mu_k x}$$

converges uniformly to zero. The coefficients are $o(k^{-2})$. By Lemma 10, there is a real number t such that

$$c_k/\mu_k^2 + m/p \cdot \lambda_k/\mu_k^2 \cos \pi\mu_k + l/p \cdot \lambda_k/(i\mu_k^3) \cdot \cos \pi\mu_k = ti(\lambda_k/\mu_k) \cdot \cos \pi\mu_k.$$

Since $c_k = o(1)$, it follows from (90) that $t = 0$. This proves the theorem.

Appendix.

17. The simple arguments of this paper, depend on the fact that the integrals

$$\int_{-\pi}^{\pi} \cos \mu_\nu x \cos \mu_n x \, dx, \qquad \int_{-\pi}^{\pi} \sin \mu_\nu x \sin \mu_n x \, dx,$$

where ν, n are different positive integers, can be expressed as a product of a function of μ_ν and a function of μ_n by means of the equation (1) which μ_ν and μ_n satisfy. The question arises: For what relatively prime polynomials A, B with real coefficients, one of which is even and the other odd, have the above integrals the form described?

For $\nu \neq n$ we have

$$\begin{aligned}\int_{-\pi}^{\pi} \cos \mu_\nu x \cos \mu_n x \, dx &= \sin \mu_\nu \pi \cos \mu_n \pi (2\mu_\nu/\mu_\nu^2 - \mu_n^2) \\ &\quad + \sin \mu_n \pi \cos \mu_\nu \pi (2\mu_n/\mu_n^2 - \mu_\nu^2) \\ &= 2 \cos \mu_\nu \pi \cos \mu_n \pi/(\mu_\nu^2 - \mu_n^2) [\mu_\nu B(\mu_\nu)/A(\mu_\nu) - \mu_n B(\mu_n)/A(\mu_n)],\end{aligned}$$

by (1). We therefore require that the value of the symmetric polynomial

$$P(x, y) = (yB(y)A(x) - xB(x)A(y))/(x^2 - y^2) \tag{97}$$

shall be of the form $b_\nu b_n$ when $x = \mu_\nu$, $y = \mu_n$, $n \neq \nu$. We shall prove (see Lemma 11 below), that this can only happen if $P(x, y) = f(x)f(y)$, where $f(x)$ is a polynomial.

Again,

$$\int_{-\pi}^{\pi} \sin \mu_\nu x \sin \mu_n x \, dx$$
$$= 2 \cos \mu_\nu \pi \cos \mu_n \pi / (\mu_\nu^2 - \mu_n^2) [\mu_n B(\mu_\nu)/A(\mu_\nu) - \mu_\nu B(\mu_n)/A(\mu_n)].$$

We therefore require that the value of the symmetric polynomial

$$Q(x,y) = (xA(x)B(y) - yA(y)B(x))/(x^2 - y^2) \tag{98}$$

shall be of the form $c_\nu c_n$ for $x = \mu_\nu$, $y = \mu_n$, $\nu \neq n$. According to Lemma 11, this can only happen if $Q(x,y) = g(x)g(y)$, where $g(x)$ is a polynomial.

There are two possibilities. (i) A even, B odd. Put $x = 0$ in P. Then $A(0)B(y)/y = -f(0)f(y)$. Now $A(0) \neq 0$, since $(A,B) = 1$. Then $f(0) \neq 0$. Hence

$$yB(y)A(x) - xB(x)A(y) = (y^2 - x^2)f(x)/f(0) \cdot B(y)/y \cdot A(0). \tag{99}$$

Hence $B(y)/y$ divides $A(y)$. Since $(A,B) = 1$, $B(y) = by$ with a constant b. By (99), A is a quadratic. Suppose that its leading coefficient is 1. Then there are two cases, $A(y) = y^2 + a$, $A(y) = 1$.

Consider next, the possibility (ii), A odd, B even. We put $x = 0$ in Q and find that $A(y) = ay$ with a constant a, and that B is a quadratic. If its leading coefficient is 1, then there are two cases, $B(y) = y^2 + b$, $B(y) = 1$.

Thus, the only cases in which the integrals are of the required form, are the two cases considered in this paper, and the two cases for which (1) becomes

$$(\alpha) \qquad \sin \pi z - bz \cos \pi z = 0, \qquad (\beta) \qquad az \sin \pi z - \cos \pi z = 0,$$

respectively. The case (α) has been the subject of recent investigations (see the remarks in [4], 324). The case (β) can be treated on similar lines.

We now turn to the lemma mentioned above. We note first, that when A is even and B is odd, $P(x,y)$ and $Q(x,y)/(xy)$ are symmetric polynomials in $\xi\ (= x^2)$ and $\eta\ (= y^2)$; and that when A is odd and B is even, the same is true of $P(x,y)/(xy)$ and $Q(x,y)$. It therefore suffices to prove

LEMMA 11. *If $p(\xi,\eta)$ is a symmetric polynomial, not identically zero, and*

$$p(a_r, a_s) = b_r b_s, \qquad (r \neq s), \tag{100}$$

where $0 < a_1 < a_2 < \cdots$, then $p(\xi,\eta) = c(\xi)c(\eta)$, where $c(\xi)$ is a polynomial in ξ.

Proof. If, for some r, $b_r = 0$, then $p(a_r, a_s) = 0$, $(s \neq r)$. Hence, $p(a_r, \eta) \equiv 0$. Let $(\xi - a_r)^n$ be the highest power of $(\xi - a_r)$ which divides $p(\xi, \eta)$. We consider $p(\xi, \eta)[(\xi - a_r)(\eta - a_r)]^{-n}$, and omit a_r from the sequence $a_1, a_2, \cdots$. Thus, we may suppose that $b_r \neq 0$ $(r = 1, 2, \cdots)$.

Then, if $r \neq s$, $p(a_t, a_r)/b_r = p(a_t, a_s)/b_s$ for $t \neq r$, $t \neq s$. Hence

$$p(\xi, a_r)/b_r = p(\xi, a_s)/b_s \tag{101}$$

for all ξ. Write

$$c(\xi) = p(\xi, a_1)/b_1. \tag{102}$$

Then by (101), on putting $s = 1$,

$$p(\xi, a_r) = b_r c(\xi) \tag{103}$$

for $r = 2, 3, \cdots$, and also for $r = 1$, by (102).

For a given s, choose an $r \neq s$ in (103). Put $\xi = a_s$. By (100), $c(a_s) = b_s$ for $s = 1, 2, \cdots$. Thus, by (103),

$$p(\xi, a_r) = c(a_r)c(\xi), \qquad (r = 1, 2, \cdots), \tag{104}$$

and so $p(\xi, \eta) = c(\xi)c(\eta)$.

18. *Historical note.* The investigations of this paper are based on [3], Theorem 1. There are points of similarity between the proof of that theorem and the investigations of Dini [1], who, writing before the era of the Lebesgue integral, could not easily attain to general and precise results. In a subsequent paper [2], Dini considered the problem of deciding when the equation (1) has all its roots real and simple. He examined the equations which form the subject of [3], Theorems 2, 3, 4; but his results are fragmentary.

19. *(Added 1 May, 1955).*

The series we have considered in this paper, can be written in the form

$$\int_{-\infty}^{\infty} e^{ixu}\, d\phi(u) = \lim_{\omega \to \infty} \int_{-\omega}^{\omega} e^{ixu} d\phi(u), \tag{105}$$

where $\phi(u)$ is constant in each open interval (μ_k, μ_{k+1}), $k = 0, \pm 1, \pm 2, \cdots$, and $\phi(\mu_k +) - \phi(\mu_k -) = c_k$. A theory of integrals (105) has been developed ([5], [7]) partly with the object of enabling one to deduce analogues of theorems concerning ordinary trigonometric series

$$\sum c_n e^{inx} \qquad c_{-n} = \bar{c}_n \tag{106}$$

without having to write out analogous proofs. As pointed out by the referee, this theory has a direct bearing on some of our results, in particular, Theorems 9 and 12.

Consider for example, the following proposition which generalizes Theorem 9: *If*

$$\sum' c_k e^{i\mu_k x} \qquad c_{-k} = \bar{c}_k \tag{107}$$

converges to $f(x)$ *for* $|x| < \pi$, *where* $f(x)$ *is finite and* $\varepsilon L(-\pi, \pi)$, *then there is a real constant* c *such that*

$$c_k = ci\mu_k\gamma_k + g_0\gamma_k/(1-\gamma_0) + g_k, \qquad (k = \pm 1, \pm 2, \cdots),$$

where $g_k = \lambda_k \int_{-\pi}^{\pi} f(t) e^{-i\mu_k t}\, dt$, $(k = 0, \pm 1, \pm 2, \cdots)$.

To prove this, we observe that the hypothesis implies that $c_k = o(1)$. Let ϵ be a small positive number. By [7], Theorem 4, there is a series

$$\sum \omega_n e^{inx} \qquad \omega_{-n} = \bar{\omega}_n, \omega_n = o(1), \tag{108}$$

such that (107), (108) are uniformly equiconvergent in $(-\pi + \epsilon, \pi - \epsilon)$, i.e.,

$$\sum_{|k| \leqq n}' c_k e^{i\mu_k x} - \sum_{|k| \leqq n} \omega_k e^{ikx}$$

converges uniformly to zero. Since (108) converges to $f(x)$ in $(-\pi + \epsilon, \pi - \epsilon)$, the theory of trigonometric series ([6], 286(i)) tells us that (108) is uniformly equiconvergent with the Fourier series S of a function of period 2π which equals $f(x)$ in $(-\pi + 2\epsilon, \pi - 2\epsilon)$. Let

$$f(x) \sim \sum g_k e^{i\mu_k x}, \qquad (-\pi, \pi). \tag{109}$$

By Lemma 1, the series (109) is uniformly equiconvergent with the Fourier series of $f(x)$ in $(-\pi, \pi)$. Hence the series (109) is uniformly equiconvergent with S in $(-\pi + 3\epsilon, \pi - 3\epsilon)$, and hence with (108) and with (107). Since ϵ is arbitrary, it follows that

$$\sum_{|k| \leqq n}' c_k e^{i\mu_k x} - \sum_{|k| \leqq n} g_k e^{i\mu_k x}$$

converges to zero for $|x| < \pi$.

Using the notation and results of Section 11, $\sum' g_0\gamma_k e^{i\mu_k x}/(1-\gamma_0)$ converges uniformly to g_0 in $(-\pi, \pi)$. Hence

$$\sum' [c_k - g_k - g_0\gamma_k/(1-\gamma_0)]\, e^{i\mu_k x}$$

converges to zero in $(-\pi, \pi)$. By Theorem 9, the desired result follows.

REFERENCES.

[1] U. Dini, *Serie di Fourier e altre rappresentazione analitiche delle funzioni di una variabile reale*, Pisa, 1880, pp. 174-211.

[2] ———, "Sugli sviluppi in serie $\frac{1}{2}a_0 + \sum(a_n \cos \lambda_n x + b_n \operatorname{sen} \lambda_n x)$ dove le λ_n sono radici della equazione trascendente $F(z)\cos \pi z + F_1(z)\operatorname{sen} \pi z = 0$," *Annali di Matematica*, ser. 3, vol. 26 (1917), pp. 261-284.

[3] S. Verblunsky, "On the roots of a transcendental equation which occurs in the theory of trigonometric series," *Mathematische Zeitschrift*, vol. 61 (1954), pp. 324-335.

[4] ———, "Some theorems on F. A. series," *Rendiconti del Circolo Matematico di Palermo*, ser. 2, vol. 3 (1954), pp. 89-105.

[5] F. Wolf, "Contributions to a theory of summability of trigonometric integrals," *University of California Publications in Mathematics*, New Series, vol. 1 (1947), pp. 159-228.

[6] A. Zygmund, *Trigonometrical Series*, Warsaw, 1935.

[7] ———, "On trigonometric integrals," *Annals of Mathematics*, vol. 48 (1947), pp. 393-440.

ISOMORPHISM BETWEEN H^p AND L^p.*

By R. P. Boas, Jr.

Let p be finite and greater than 1. By L^p we understand the Banach space of complex-valued measurable functions f (those differing only on a set of measure zero being identified) of period 2π, such that $\int_0^{2\pi} |f(x)|^p\,dx$ is finite, with norm

$$\|f\| = \{\int_0^{2\pi} |f(x)|^p\,dx\}^{1/p}.$$

An element f of L^p has a formal Fourier series,

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}, \qquad c_n = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}\,dx. \tag{1}$$

M. Riesz's theorem on conjugate functions may be stated in the following form (cf., in particular, [1], p. 791):

If $f \in L^p$ has the Fourier series (1), *then* (1) *with c_n replaced by* 0 *for $n < 0$ is the Fourier series of a function f_+ of L^p and $\|f_+\| \leqq A_p\|f\|$, with A_p independent of f.*

By H^p we understand the Banach space, with the L^p norm, whose elements are the functions f_+ obtained from L^p-functions according to this theorem. It is known ([3]; [2], p. 293) that H^p can alternatively be defined as the Banach space (with the natural norm) of boundary values of functions F which are analytic in $|z| < 1$ and satisfy

$$\int_0^{2\pi} |F(re^{i\theta})|^p\,d\theta \leqq M, \qquad 0 < r < 1,$$

with M independent of r.

Our object is to give a simple proof of the fact, which may be useful in investigations involving the spaces H^p, that *the topological vector-spaces H^p and L^p are isomorphic.* The question of whether there is an isomorphism

* Received March 7, 1955.

between them which preserves the metric is left open, as well as the question of the relationships of H^1 and H^∞ to L^1 and L^∞.

Let f be a given element of L^p with Fourier series (1). Let f_+ be defined as in Riesz's theorem and let $f_- = f - f_+$; both these functions have period 2π. Map f into the function g defined by

$$g(x) = f_+(2x) + e^{-ix} f_-(-2x);$$

then the Fourier series of $g(x)$ is $c_0 + \sum_{n=1}^{\infty} (c_n e^{2inx} + c_{-n} e^{(2n-1)ix})$, so that $g \in H^p$ and $\| g \| \leqq 2A_p \| f \|$. Thus the mapping is bounded; and it is clearly distributive. It is also one-to-one and onto. For, let g be a given element of H^p with Fourier series $\sum_{n=0}^{\infty} b_n e^{inx}$. The series

$$\sum_{n=0}^{\infty} b_{2n} e^{2inx}, \qquad \sum_{n=1}^{\infty} b_{2n-1} e^{(2n-1)ix} \tag{2}$$

are Fourier series of functions in H^p; indeed, the first is the Fourier series of $g_1(x) = \frac{1}{2}\{g(x) + g(x+\pi)\}$, and the second is then the Fourier series of $g_2(x)e^{-ix}$, where $g_2(x) \sim \sum_{n=1}^{\infty} b_{2n-1} e^{2inx}$. We can map g into the element f of L^p given by $f(x) = g_1(x/2) + g_2(-x/2)$. Applied to the element g obtained from f by the original mapping, this leads back to f; so the inverse mapping from H^p to L^p is one-to-one. Finally, this inverse mapping is bounded, either by another application of Riesz's theorem or because it is the inverse of a one-to-one bounded mapping.

NORTHWESTERN UNIVERSITY.

REFERENCES.

[1] S. Bochner, "Additive set functions on groups," *Annals of Mathematics*, vol. 40 (1939), pp. 769-799.

[2] W. W. Rogosinski and H. S. Shapiro, "On certain extremum problems for analytic functions," *Acta Mathematica*, vol. 90 (1953), pp. 287-318.

[3] A. E. Taylor, "On certain Banach spaces whose elements are analytic functions," *Actas Acad. Ci. Lima*, vol. 12 (1949), pp. 31-43.

ON FINITE GROUPS WITH CYCLIC SYLOW SUBGROUPS FOR ALL ODD PRIMES.*

By MICHIO SUZUKI.

The purpose of this paper is to determine the structure of some finite groups in which all Sylow subgroups of odd order are cyclic. This assumption on Sylow subgroups simplifies the structure of groups considerably, but the structure of 2-Sylow subgroups might be too complicated to make any definite statement on the structure of the groups. In this paper, therefore, we shall make another assumption on 2-Sylow subgroups, and our main result may be stated as follows.

Let G be a non-solvable group of finite order. We assume that all Sylow subgroups of odd order are cyclic, and moreover that a 2-Sylow subgroup is either (a) a dihedral group, or (b) a generalized quaternion group. Then G contains a normal subgroup G_1 such that $[G: G_1] \leqq 2$ and $G_1 = Z \times L$, where Z is a solvable group whose Sylow subgroups are all cyclic, and L is isomorphic with the linear fractional group $LF(2, p)$ over the prime field of characteristic p in the case (a), and with the special linear group $SL(2, p)$ in the case (b).

It has been proved by Frobenius and Burnside (see Burnside [7]) that if all Sylow subgroups of a finite group G are cyclic, then G is solvable. In fact, such a group must be meta-cyclic. Zassenhaus [13] has studied the structure of such groups completely (see Zassenhaus [13], or [14], § 4). In the same paper, Zassenhaus studied the *solvable* groups in which all Sylow subgroups of odd order are cyclic and a 2-Sylow subgroup contains a cyclic subgroup of index 2, and he classified all possible types of such groups. His results are used frequently in the course of the proof of our main result. We refer for these results to the paper [13] of Zassenhaus.

In the first section we shall give a short summary of known properties of linear groups, which are necessary for our purpose, referring mostly to the work by Schur [10]. Some general lemmas on group characters are proved in the next section 2. Our Lemmas 4 and 5 may be regarded as slight generalizations of a result by Brauer and they are main tools for the proof. Section 3 contains several characterizations of linear groups by their group

* Received November 11, 1954.

theoretical properties, obtained by Brauer, Suzuki, and Wall. Sections 4 and 5 are devoted to the proof of our main theorem in case of dihedral 2-Sylow subgroups. The case when 2-Sylow subgroups are generalized quaternion groups is treated in Section 6, where we consider a slightly more general situation (see Theorem D). In the proof of Theorem D we shall use very delicate properties of group characters, namely the block theory of Brauer. The final section contains remarks on the applications and the converse of our main theorem.

We use the standard notations. The *linear fractional group* $LF(2,p)$ is the totality of linear fractional transformations $x' = (\alpha x + \beta)/(\gamma x + \delta)$ with elements α, β, γ, δ in the prime field F of characteristic p, and with determinant $\alpha\delta - \beta\gamma = 1$. The *special linear group* $SL(2,p)$ is the totality of non-singular 2×2 matrices over F with determinant 1. A *generalized quaternion group* is a non-cyclic 2-group which contains only one subgroup of order 2. It is generated by two elements σ and ρ with the relations

$$\sigma^{2^{n+1}} = 1, \quad \rho^2 = \sigma^{2^n} \quad \text{and} \quad \rho\sigma\rho^{-1} = \sigma^{-1}.$$

The central factor group of a generalized quaternion group is a *dihedral group*. A finite group is called a *Z-group* if all Sylow subgroups are cyclic. For the properties of Z-groups see Zassenhaus [13]. Group elements are denoted by Greek letters σ, ρ and τ; τ stands usually, not exclusively, for elements of order 2. An element σ is called *p-regular* if the order of σ is prime to p. Otherwise σ is *p-singular*.

1. We shall make preliminary remarks on the special linear group $SL(2,p)$ and the linear fractional group $LF(2,p)$ over the prime field of characteristic p.

The group $SL(2,p)$ contains only one element of order 2, which generates the center, and the central factor group is isomorphic with $LF(2,p)$. It has been proved by Schur, [10], p. 120, that if a finite group G contains a normal subgroup Z of order 2, and if the factor group G/Z is isomorphic with $LF(2,p)$, then G is either isomorphic with $SL(2,p)$, or is a direct product of Z and $LF(2,p)$. In particular we have

LEMMA 1. *If a finite group G contains only one subgroup Z of order 2, and if the factor group G/Z is isomorphic with $LF(2,p)$, then G is isomorphic with $SL(2,p)$.*

We shall consider the group of all automorphisms of $LF(2,p)$. The following lemma is proved.

LEMMA 2. *The index of the group of inner automorphisms in the group of all automorphisms of* $LF(2,p)$ *is two.*

Proof. Let L be the totality of linear fractional transformations over the $GF(p)$. The transformations with determinant 1 form a normal subgroup of index 2, isomorphic with $LF(2,p)$. Now let A be the group of all automorphisms of $LF(2,p)$. Then A contains a subgroup isomorphic with L. $LF(2,p)$ may be regarded as a subgroup I of A, consisting of all inner automorphisms. I contains $p+1$ subgroups of order p, and A may be regarded as a substitution group on these subgroups. Since $A \supseteq L$, A is at least triply transitive. Suppose that an element σ leaves invariant three subgroups P_1, P_2 and P_3. Consider the normalizer N_i of P_i in I. Since σ leaves invariant P_i, every N_i is also fixed by σ, and so are $N_1 \cap N_2$ and $N_1 \cap N_3$. Now N_1 may be generated by two elements τ and ρ such that

$$\{\tau\} = P_1, \qquad \{\rho\} = N_1 \cap N_2, \qquad \{\tau\rho\} = N_1 \cap N_3$$

and

$$\rho\tau\rho^{-1} = \tau^r.$$

Since σ maps $\{\tau\}$, $\{\rho\}$ and $\{\tau\rho\}$ into themselves, we have

$$\sigma\tau\sigma^{-1} = \tau^m, \quad \sigma\rho\sigma^{-1} = \rho^n \quad \text{and} \quad \sigma(\tau\rho)\sigma^{-1} = (\tau\sigma)^k,$$

where m, n and k are positive integers. These equations imply that $\rho^n\tau^m\rho^{-n} = \tau^{mr}$ or $mr \equiv mr^n \pmod{p}$. Since $r^{(p-1)/2} \equiv 1 \pmod{p}$, we have $n \equiv 1 \pmod{(p-1)/2}$. Moreover since $k \equiv n \pmod{(p-1)/2}$, we may assume that $n = k = 1$, and hence $m = 1$. Thus σ induces the identical automorphism on N_1, and similarly on N_2. Hence σ transforms all elements of I into themselves, and so is in the centralizer of I in A. On the other hand, the centralizer of I in A contains only the identity, since the center of I is trivial. This proves $\sigma = 1$. Hence A is representable as a triply transitive permutation group on $p+1$ letters such that no element except the identity leaves invariant three letters. By a theorem of Zassenhaus [12], A is isomorphic with a linear fractional group, and so A coincides with L. This proves Lemma 2.

We shall now consider the irreducible characters of $SL(2,p)$. The irreducible characters of $SL(2,p)$ have been determined by Schur [10] (see in particular the table on p. 128) and independently by Jordan [9].

Let G be the special linear group $SL(2,p)$ over the prime field of characteristic p. We are interested in the structure of 2-blocks of characters of G. For the theory of blocks see [1], [2] and [3]. Let ϵ be 1 or -1,

so that $p \equiv \epsilon \pmod 4$, and let 2^α be the highest power of 2 which divides $p-\epsilon$:

$$p-\epsilon = 2^\alpha h, \qquad (h,2)=1 \qquad (\alpha \geqq 2).$$

If we set $(p+\epsilon)/2 = 2a+1$, G has exactly a 2-blocks of defect 1. This may be proved either directly from the table of characters, or by using a result of Brauer [3], § 8, since the factor group G/Z modulo the center Z of G is isomorphic with $LF(2,p)$, and $LF(2,p)$ has exactly a 2-blocks of defect 0. From the general theory of Brauer [1], each 2-block of defect 1 consists of two characters, and they take the same value on every 2-regular element. The degree of characters in a block of defect 1 is $p-\epsilon$. Some characters of degree $p+\epsilon$ may belong to blocks of defect α. In fact, if we put $h-1=2b$, G has exactly b 2-blocks of defect α. From the table of characters we see that each block of defect α contains 2^α characters which take the same values on every 2-regular element.

The first 2-block B_1 of G is the only block of defect $\alpha+1$. It consists of the principal character 1, one character ϕ_1 of degree p, two p-conjugate characters ϕ_2 and ϕ_3 of degree $(p+\epsilon)/2$, which belong to the representations of the central factor group, two conjugate characters ϕ_4 and ϕ_5 of degree $(p-\epsilon)/2$, and $2^{\alpha-1}-1$ characters ϕ_i $(i=6,7,\cdots,2^{\alpha-1}+4)$ of degree $p+\epsilon$.

There are several linear combinations of these characters which vanish on all 2-regular elements. The following table may be obtained from the table of characters (using suitable notations).

Relations \ order	prime to 2	$p+\epsilon$	2^α	2
$1+\epsilon\phi_1-\epsilon\phi_2-\epsilon\phi_3$	0	0	4	0
$1-\epsilon\phi_2+\epsilon\phi_4$	0	2	2	$1-\epsilon p$
$1-\epsilon\phi_3+\epsilon\phi_5$	0	2	2	$1-\epsilon p$
$\phi_i-\phi_2-\phi_3$	0	0	$2\epsilon+\epsilon\theta_i$	$\begin{matrix}0\\ -2(p+\epsilon)\end{matrix}$

TABLE I

In the last line of Table I, θ_i $(i=6,7,\cdots,2^{\alpha-1}+4)$ are the values of non-linear characters of a generalized quaternion group of order $2^{\alpha+1}$, taken in a suitable arrangement; and, in the lower right hand corner, the value 0 is for those ϕ_i which belong to representations of the central factor group and the value $-2(p+\epsilon)$ is for those which do not.

2. In this section we shall prove some lemmas on group characters. Some of these lemmas have been obtained by Brauer and Fowler [5].

Let $C_1, C_2, \cdots, C_k$ be conjugate classes of a finite group G of order g, and $s_1, \cdots, s_k$ the totality of distinct irreducible characters. The product C_iC_j may be written as a sum of conjugate classes: $C_iC_j = \sum_{l=1}^{k} a_{ijl}C_l$, where a_{ijl} are non-negative integers. The following formula is a consequence of orthogonality relations (see [5]):

$$(*) \qquad a_{ijl} = g_i g_j g^{-1} \sum_{\mu} z_\mu^{-1} s_\mu(C_i) s_\mu(C_j) \bar{s}_\mu(C_l).$$

In this formula, g_i is the number of elements in C_i, z_μ is the degree of s_μ and $\bar{s}_\mu$ indicates the conjugate complex of s_μ. In the application to our case, classes C_i and C_j contain elements of order 2. We have

LEMMA 3. *Let G be a finite group of even order. Assume that G has only one class T which contains an element of order 2, and that $T^2 = \sum_{i=1}^{k} a_iC_i$. Then we have*

$$a_l = gn^{-2} \sum_{\mu} z_\mu^{-1} s_\mu(T)^2 s_\mu(C_l)$$

where n is the order of the normalizer of an element of order 2. Moreover if the class C_i is not self-reciprocal, then $a_i = 0$.

Proof. Let σ and τ be two elements of order 2. Then the inverse of $\sigma\tau$ is $\tau\sigma$, which is obviously conjugate to $\sigma\tau$. This proves the last assertion. The first assertion is a consequence of (*).

By a (generalized) character, we shall mean a linear combination of irreducible characters with rational integral coefficients. Let ϕ be a character of G. Then $\phi = \sum_\mu a_\mu s_\mu$, and all a_μ are rational integers, uniquely determined by ϕ; namely $a_\mu = \int_G \phi\bar{s}_\mu$. (The symbol $\int_G$ means the average on G, i.e. the summation over all elements of G divided by the order of G.) The sum of the square of these coefficients is denoted by $w(\phi)$, and may be calculated as $w(\phi) = \sum_\mu a_\mu^2 = \int_G |\phi|^2$, where $|\phi|$ is the absolute value of ϕ.

We consider a subgroup H of G and a character ψ of H. The induced character ψ^* of G may be defined by

$$\psi^*(\sigma) = h^{-1} \sum_{\rho \in G} \psi(\rho\sigma\rho^{-1}),$$

where h is the order of H, and we have put $\psi(\sigma) = 0$ if $\sigma \notin H$. Clearly, ψ^*

is a character of G (in our sense). If $\psi = \sum a_i\psi_i$, then we have $\psi^* = \sum a_i\psi_i^*$.

A conjugate class C of a subgroup H is called *special* (with respect to G), if C contains an element σ such that the centralizer of σ in G is contained in H. We shall prove the following lemma.

LEMMA 4. *Let $C_1, C_2, \cdots, C_r$ be special classes of H, and $\sigma_i \in C_i$ $(i = 1, 2, \cdots, r)$. Assume that σ_i is not conjugate to σ_j even in G if $i \neq j$. If ψ is a character of H which vanishes outside of these C_i, then $\psi^*(\sigma_i) = \psi(\sigma_i)$ $(i = 1, 2, \cdots, r)$, and $\psi^*(\sigma) = 0$ if σ is not conjugate to any σ_i. Moreover we have $w(\psi^*) = w(\psi)$.*

Proof. By the definition of induced character,

$$\psi^*(\rho) = h^{-1} \sum_{\tau \in G} \psi(\tau\rho\tau^{-1}).$$

Hence if $\psi^*(\rho) \neq 0$, at least one term of the right hand side does not vanish. But ψ is zero outside of C_i, and so $\psi^*(\rho) \neq 0$ implies that ρ is conjugate to some σ_k. This proves the second assertion.

If $\rho = \sigma_i$, then $\psi(\tau\rho\tau^{-1}) \neq 0$ only when $\tau\rho\tau^{-1} \sim \sigma_k$ for some k. By our assumption σ_i is conjugate to σ_k only when $k = i$. Hence $\psi(\tau\rho\tau^{-1}) \neq 0$ implies that $\tau\rho\tau^{-1}$ is conjugate to σ_i in H, i.e. $\tau\rho\tau^{-1} = \nu\sigma_i\nu^{-1}$ $(\nu \in H)$. Then $\nu^{-1}\tau$ is in the normalizer of σ_i. Since C_i is special, the normalizer of σ_i is contained in H. Hence we conclude that $\tau \in H$. We see now that

$$\psi^*(\sigma_i) = h^{-1} \sum_{\tau \in H} \psi(\tau\sigma_i\tau) = \psi(\sigma_i).$$

Let N_i be the normalizer of σ_i in G. By our assumption, we may take suitable σ_i so that N_i is a subgroup of H. The number of elements which are conjugate to σ_i in G is denoted by g_i, and similarly h_i is the number of elements in C_i. Then $g_i = [G : N_i]$ and $h_i = [H : N_i]$. Hence we have $g_i = [G : H]h_i$ $(i = 1, 2, \cdots, r)$. We shall now compute $w(\psi^*)$. We have

$$\begin{aligned} w(\psi^*) &= \int_G |\psi^*|^2 = g^{-1} \sum_i g_i |\psi^*(\sigma_i)|^2 \\ &= g^{-1} \sum_i [G : H]h_i |\psi(\sigma_i)|^2 = h^{-1} \sum_i h_i |\psi(\sigma_i)|^2 = \int_H |\psi|^2 = w(\psi), \end{aligned}$$

since $\psi = 0$ outside of C_i, which proves the statement.

As an application we shall prove the following lemma, due essentially to Brauer (see [4]).

LEMMA 5. *Let $C_1, \cdots, C_r$ be special classes of H satisfying the same assumption as in Lemma 4. If H has s irreducible characters $\phi_1, \cdots, \phi_s$ such that $\phi_i = \phi_j$ outside of C_k, and $\phi_i(\sigma) = 0$ for any element σ of H which*

is conjugate to σ_k in G but not in H, and if $s \geqq 2$, then G has exactly s irreducible characters $s_1, \cdots, s_s$ such that

$$\phi_i^* = \epsilon s_i + a \sum_{k=1}^{s} s_k + \Delta, \qquad (i = 1, 2, \cdots, s),$$

where $\epsilon = \pm 1$, $a \geqq 0$, $a + \epsilon \geqq 0$ and Δ is a character of G which is independent of i and does not contain any s_j. All characters of G other than these s_i contain each ϕ_k with the same multiplicity (as characters of H).

Definition. These characters $s_1, \cdots, s_s$ are called the *exceptional* characters associated with $\phi_1, \cdots, \phi_s$.

Proof. Since $s \geqq 2$, we may consider $s-1$ characters $\psi_i = \phi_i - \phi_s$ $(i = 1, 2, \cdots, s-1)$ of H which satisfy the assumption of Lemma 4. Hence $w(\psi_i^*) = w(\psi_i) = 2$, and so each induced character ψ_i^* contains exactly 2 different irreducible characters of G with multiplicity $+1$ or -1. Since

$$w(\psi_i^* - \psi_j^*) = w(\psi_i - \psi_j) = 2 \qquad (i \neq j),$$

and since $\psi_i^* = 0$ for the identity of G, we may write $\psi_i^* = \epsilon(s_i - s_s)$, $(i = 1, 2, \cdots, s-1)$, with $\epsilon = +1$ or -1. Hence we have $\phi_i^* - \epsilon s_i = \phi_s^* - \epsilon s_s$. This implies that ϕ_i^* may be decomposed in the following manner:

$$\phi_i^* = \epsilon s_i + \sum_{j=1}^{s} a_j s_j + \Delta,$$

where Δ is either identically 0 or a sum of irreducible characters of G other than $s_1, \cdots, s_s$. Consider now the difference $w(\phi_i^*) - w(\phi_s^*)$. This is equal to

$$g^{-1} \sum_{\sigma \in G} (\phi_i^*(\sigma)\phi_i^*(\sigma^{-1}) - \phi_s^*(\sigma)\phi_s^*(\sigma^{-1})).$$

If σ is not conjugate to any σ_k $(k = 1, 2, \cdots, r)$, then $\phi_i^*(\sigma) = \phi_s^*(s)$. If σ is conjugate to σ_k, then from our assumption on ϕ_i we see that $\phi_i^*(\sigma_k) = \phi_i(\sigma_k)$. Hence as in the proof of Lemma 4, we conclude that

$$w(\phi_i^*) - w(\phi_s^*) = w(\phi_i) - w(\phi_s) = 0.$$

Hence we have $(a_i + \epsilon)^2 + a_s^2 = a_i^2 + (a_s + \epsilon)^2$, or $a_i = a_s$. Hence we obtain $\phi_i^* = \epsilon s_i + a \sum_{k=1}^{s} s_k + \Delta$, as claimed.

3. In the proof of our main theorems, we shall need some group theoretical characterization of $LF(2, p)$. We shall state them in the following.

LEMMA 6. *Let G be a finite simple group of even order, satisfying the following property: If two cyclic subgroups U and V have a non-trivial intersection, i.e., $U \cap V \neq e$, and if at least one of these subgroups has an even order, then $U \cup V$ is cyclic. Then G is isomorphic with $LF(2, q)$ for some prime power q.*

This result is due to Brauer. See [6]. From this lemma, we conclude the following

LEMMA 7. *Let G be a simple group of finite even order. We assume that all Sylow subgroups of odd order are cyclic, and that the normalizer of an element of order 2 is of dihedral type. Then G is isomorphic with $LF(2, p)$ for some prime p.*

Proof. Let τ be an element of order 2, of which the normalizer N is of dihedral type. Put $T = \{\tau\}$. Consider a 2-Sylow subgroup S of N. S is contained in a 2-Sylow subgroup S^* of G. If $S \neq S^*$, T is not a characteristic subgroup of S, and this implies that S is of order 4. Since S is the normalizer of T in S^*, S^* is generated by two elements σ and τ such that $\sigma^{2^n} = 1$, $\tau\sigma\tau^{-1} = \sigma^r$, where $r = -1$ or $-1 + 2^{n-1}$. (See Lemma 4 of [11]). Hence τ is not conjugate to the central element of S^* in G. By a theorem of Grün [8], G is not perfect. This contradicts our assumption. Hence $S = S^*$ is a 2-Sylow subgroup of G. We see now that the totality of elements of order 2 forms a single conjugate class of G. This follows also from the theorem of Grün (loc. cit.).

Now N contains a normal subgroup Z such that $ZS = N$ and $Z \cap S = e$. If $Z = e$, then G satisfies the assumption of Lemma 6. Hence our assertion follows immediately.

Assume that $Z \neq e$. Consider any subgroup U of Z. The centralizer V of U in G has a cyclic 2-Sylow subgroup. Hence V is a Z-group. We shall prove that $V \subseteq N$ using induction on the index $[Z : U]$. First of all assume that $U = Z$. Suppose $V \not\subseteq N$. Denoting by M the maximal cyclic subgroup of N with index 2, V contains a normal subgroup W such that $WM = V$ and $W \cap M = Z$. In the factor group V/Z, the normalizer of TZ/Z is M/Z. Hence the element τ of T induces an automorphism f in W/Z which maps every element onto its inverse. Now W is a normal subgroup of the normalizer of Z in G, which contains an element σ of order 2, $\sigma \notin T$. This element σ induces an automorphism of each Sylow subgroup of W/Z, which is either the identical mapping or the same one as f. This implies that every prime divisor of $[W : Z]$ also divides the order of Z, since

any element of order 2 is conjugate to τ. On the other hand the order of the commutator subgroup of a Z-group is prime to the index of commutator group, which is a multiple of the order of its center. Hence the commutator subgroup of V must be e, and so V is cyclic. Since $V \supseteq T$, and we get a contradiction. Hence the centralizer of Z is contained in N.

Now we consider any subgroup U of Z. Consider all subgroups U_i of Z such that $U_i \supset U$ and U_i/U are Sylow subgroups of Z/U. By the inductive hypothesis, each U_i commutes only with elements of N, since the centralizer of U_i is M, and so the normalizer of U_i is contained in the normalizer N of T. Suppose that the centralizer V of U is not contained in N. In V/U, each U_i/U is a Sylow subgroup and is contained in the center of its normalizer. Hence by a theorem of Burnside (see e.g. [14], p. 133), V contains a normal subgroup W such that $WM = V$ and $W \cap M = U$. We may apply the same argument as before to get a similar contradiction.

We have thus proved that the centralizer of any subgroup of Z coincides with M. Hence our group G satisfies the property of Lemma 6, and is isomorphic with $LF(2, p)$ for some prime number p. This proves Lemma 7.

Finally we shall state the following result (see [6]).

LEMMA 8. *Let G be a non-solvable group of finite even order. If G contains an abelian subgroup A of even order such that $A \neq e$ and the centralizer of any element $\neq 1$ of A is contained in A, then G is isomorphic with $LF(2, 2^n)$.*

4. This section is devoted to the proof of the following theorem:

THEOREM A. *Let G be a non-solvable group of even order. We assume that all l-Sylow subgroups for odd l are cyclic, and that 2-Sylow subgroups are of order 4. Then G is a direct product of $LF(2, p)$ and a Z-group.*

Proof. We shall use an induction on the order g of G.

First of all, suppose that the commutator subgroup G' of G is a proper subgroup of G. Then G' satisfies all assumptions of this Theorem A, and has a smaller order than G. We may, therefore, apply our inductive hypothesis to G'. G' is a direct product of a Z-group U and a subgroup L, isomorphic with $LF(2, p)$ for some p. Since U is solvable, L is a characteristic subgroup of G', and hence L is a normal subgroup of G. Consider now the centralizer V of L in G. V is a normal subgroup of G, and the factor group G/V is isomorphic with a subgroup of the group L_0 of all automorphisms of L. By Lemma 2, the index $[L_0 : L]$ is 2. On the other hand, by assump-

tion the order of G is not divisible by 8, and hence $[G:V \cup L]$ is odd. This implies that $G = V \cup L$, or $G = V \times L$.

We may therefore assume that $G = G'$, i.e. that G is perfect. Suppose that G contains a proper normal subgroup H. If $[G:H]$ were not divisible by 4, then the factor group G/H would be a Z-group, and so solvable. This is not the case. If $[G:H] \equiv 0$ (4), then by assumption H would be of odd order, and so a Z-group. Hence H would contain a characteristic subgroup of odd prime order q. This is however impossible, because no perfect group with cyclic q-Sylow subgroups can contain a normal subgroup of order q by Theorem 15 of Zassenhaus [13]. Hence G must be a simple group.

Let S_2 be a 2-Sylow subgroup of G, and let Z_2 be the centralizer and N_2 the normalizer of S_2 in G. Since G is simple, $N_2 \neq Z_2$. Hence S_2 is not cyclic, and $[N_2:Z_2] = 3$. Z_2 is a direct product of S_2 and a Z-group U: $Z_2 = U \times S_2$.

Assume that $U = e$. Take a subgroup T of order 2 in S_2. Consider the normalizer N of T in G. The factor group N/T is a Z-group, and hence N contains a normal subgroup Z such that $N = ZS_2$ and $Z \cap S_2 = e$. An element of S_2, which is not in T, induces an automorphism of Z, and this automorphism leaves only the identity invariant. Hence by a lemma of Zassenhaus ([12], Lemma 1) Z is a cyclic group and N is a dihedral group. Theorem A now follows from Lemma 7.

Suppose finally that $U \neq e$. This case will be proved impossible at the end of this section. Again denote by T a subgroup of order 2 in S_2. We consider the structure of the normalizer N of T in G.

Consider the normalizer $N(U)$ of U in G. Suppose $N(U) \neq N_2$. If the factor group $N(U)/U$ is solvable, then by Theorem 6 of Zassenhaus [13], $N(U)$ contains a normal subgroup V such that $VN_2 = N(U)$ and $V \cap N_2 = U$. Since V is of odd order, V/U is a Z-group, and hence V contains a subgroup P such that P is a normal subgroup of $N(U)$ and $[P:U] = l$ is a prime number. Then the centralizer C/U of P/U contains Z_2/U, since $N(U)/C$ is cyclic. This implies that P is a part of the centrailzer of S_2. This gives a contradiction. Hence $N(U)/U$ is not solvable. We may apply the inductive hypothesis to $N(U)$ since $G \neq N(U)$. It follows that $N(U)$ is a direct product of U and a subgroup L, which is isomorphic with some $LF(2,p)$. Let q be the smallest prime divisor of the order of U, Q the q-Sylow subgroup of U. The normalizer $N(Q)$ of Q in G contains L, and so it is not solvable. By the inductive hypothesis $N(Q) = L' \times U'$ where $L' \supseteq L$ and $U' \subseteq U$. Since q is the smallest prime divisor of the order of U, $N(Q)$ coincides with the centralizer of Q. This

implies that a q-Sylow subgroup of G is in the center of its normalizer. By a theorem of Burnside (loc. cit.), G is not simple. This contradiction comes from the assumption that $N(U) \neq N_2$, and hence $N(U) = N_2$.

Consider next a maximal subgroup M of G, which contains N_2. If M is not solvable, $N(U)$ would be larger than N_2 by the inductive hypothesis. M is, therefore, solvable. Hence by Theorem 6 of Zassenhaus [13], M contains a normal subgroup K such that $M = KN_2$ and $K \cap N_2 = U$. Since K is a Z-group, the commutator subgroup K' of K is cyclic. Take a Sylow subgroup P of K'. P is normal in M, so that the centralizer $Z(P)$ is also normal, and $M/Z(P)$ is cyclic. Hence $Z(P)K$ contains S_2, and this shows that $Z(P) \supseteq S_2$, or $P \subseteq Z_2$. Since P was any Sylow subgroup of K', we conclude that $K' \subseteq U$. This implies that U is a normal subgroup of M, or $M \subseteq N(U) = N_2$. In other words, N_2 is a maximal subgroup of G.

We shall show that U is cyclic. N_2 contains a subgroup W such that $WS_2 = N_2$ and $W \cap Z_2 = U$. Suppose that the commutator subgroup W' of W is a proper part of U. We consider any prime divisor of $[U : W']$, say q, and the q-Sylow subgroup Q of U. The normalizer $N(Q)$ of Q in G is solvable, since otherwise G would be non-simple by Burnside's theorem. Since $N_2 = U(N_2 \cap N(Q))$, $N(Q)$ has no normal subgroup of index 2. Hence again by Theorem 6 of Zassenhaus [13], $N(Q)$ contains a normal subgroup X such that $N(Q)/X$ is isomorphic with the tetrahedral group. Let Y be any Sylow subgroup of X. Then the order of the centralizer of Y in $N(Q)$ is divisible by 4, and hence some conjugate subgroup Y_0 of Y in $N(Q)$ is contained in Z_2, or $X \subseteq U$. Hence $N(Q)$ coincides with the centralizer of Q, which is impossible by virtue of Burnside's theorem. Hence we have $W' = U$, or U is cyclic.

Consider again the normalizer N of the subgroup T of order 2. Suppose that $N \subseteq N_2$. Then Z_2 is abelian, and the centralizer of any element $\neq 1$ of Z_2 coincides with Z_2. By Lemma 8 this is impossible, since we must have $U = e$ in $LF(2, 2^\mu)$.

Hence we consider the case $N \not\subseteq N_2$. We then have $N \cap N_2 = Z_2$, so that we see from Burnside's theorem (loc. cit.) that N contains a normal subgroup H such that $HS_2 = N$ and $H \cap S_2 = e$. If H' is the commutator subgroup of H, the subgroup $H' \cap U$ is a normal subgroup of both N and N_2. Since N_2 is maximal, we must have $H' \cap U = e$. H contains a cyclic subgroup X such that $X \supseteq U$, $X \cap H' = e$ and $XH' = H$. Since the normalizer of U is N_2, $X \subseteq N_2$ or $X = U$. In particular, we have $UH' = H$. The above consideration is applied to any maximal subgroup M of G which contains N. Corresponding to H, M contains a normal subgroup K such that $K \supseteq H$,

$K \cap S_2 = e$ and $KS_2 = M$. We want to show that $M = N$. If $N(H)$ is the normalizer of H in G, $N(H) = H \cup X$, where X is the normalizer of some Sylow subgroup Q of U in $N(H)$. But $N(H) \supset N_2$, so that $X = Z_2$, or $N(H) = HZ_2 = N$. Hence if $M \neq N$, the commutator subgroup K' of K is not a part of H. The element of order 2 in T induces an automorphism of $K'/K' \cap H$, which sends every element into its inverse. This implies that $[K' : K' \cap H]$ is prime to $[K' \cap H : e]$ since every Sylow subgroup of K' is cyclic. On the other hand, if q is any prime divisor of $[K' : K' \cap H]$, M contains an element of order $2q$, since the centralizer of a q-Sylow subgroup of K' is of even order. Hence $[K' : K' \cap H]$ divides the order of H, and therefore $[K' : K' \cap H]$ divides $[H : K' \cap H] = [HK' : K']$. This is a contradiction, since the order and index of K' in K must be relatively prime. We must have $M = N$; in other words, N is a maximal subgroup of G.

Let $[U : e] = u$ and $[H' : e] = v$. Then we have $g = [G : e] = 4uvm$, and u, v, m are relatively prime. By assumption $U \neq e$, or $u > 1$.

By a well-known argument of Burnside we see that two elements in H' are conjugate in G if and only if they are conjugate in N. Since N is a maximal subgroup of G and since G is simple, every conjugate class of N containing elements $\neq 1$ of H' is special with respect to G. N has $n = (v-1)/u$ non-linear irreducible characters of degree $2u$, which vanish on elements outside of TH'. Denote them by $\phi_1, \phi_2, \cdots, \phi_n$. For these characters the assumptions of Lemma 5 are satisfied. Hence we conclude that G has n "exceptional" characters $E_1, \cdots, E_n$ such that $\phi_i^* = \delta E_i + \Delta$, where $\delta = \pm 1$ and Δ is independent of i.

Similarly every conjugate class of N_2 containing elements of Z_2 with order greater than 2 is special. N_2 has $t = (u-1)/3$ characters $\theta_1, \theta_2, \cdots, \theta_t$ of degree 3, which are the characters of N_2/S_2, and u characters $\xi_1, \cdots, \xi_u$ of degree 3, which belong to representations which have kernels contained in U. All characters θ_i and ξ_j vanish outside of Z_2. Hence by Lemma 5, G has exceptional characters Θ_i and Ξ_j associated with θ_i and ξ_j respectively. (Note that $t \geqq 2$, since u is odd.) As a character of N_2 Ξ_j contains each θ_i with the same multiplicity, but Θ_k does not. Hence $\Xi_j \neq \Theta_k$.

Let $\bar{\theta}_i$ be the linear character of N whose restriction θ_i on Z_2 induces θ_i in N_2. Then we have $\theta_i^* = \bar{\theta}_i^* = \theta_i^* + \sum' \phi_k^*$, where $\sum'$ denotes the summation over all subscripts k for which ϕ_k belongs to a representation of N/T. Hence for $i \neq j$ we have

$$\bar{\theta}_i^* - \bar{\theta}_j^* = \theta_i^* - \theta_j^* = \epsilon_1 (\Theta_i - \Theta_j),$$

or $\bar{\theta}_i^* = \epsilon_1 \Theta_i + \Delta_1$, $(i = 1, 2, \cdots, t)$, where $\epsilon_1 = \pm 1$ and Δ_1 is independent of i. Similarly for the principal characters $\bar{\theta}$ of N and $\tilde{\theta}$ of Z_2, we have $\tilde{\theta}^* = \bar{\theta}^* + \sum' \phi_k^*$. We apply Lemma 4 to the character $\psi = 1 + \omega + \omega' - \theta_i$ of N_2, where $1, \omega, \omega'$ are linear characters of N_2/Z_2. Then $w(\psi^*) = w(\psi) = 4$. Since ψ^* contains the principal character of G, ψ^* must be a sum of four irreducible characters. Hence we have

$$\bar{\theta}^* - \bar{\theta}_i^* = \tilde{\theta}^* - \tilde{\theta}_i^* = 1 + \delta_1 X_1 + \delta_2 X_2 + \delta_3 X_3 \qquad (\delta_i = \pm 1).$$

One of these characters X_i must be Θ_i, and we have say $X_3 = \Theta_i$ and $\delta_3 = -\epsilon$. Hence $\bar{\theta}^* = 1 + \delta_1 X_1 + \delta_2 X_2 + \Delta_1$.

We shall consider the linear character $\bar{\xi}_j$ of N such that $\bar{\xi}_j(\tau) = -1$ for the generator τ of T, and the restriction $\tilde{\xi}_j$ of $\bar{\xi}_j$ on Z_2 induces ξ_j of N_2. By the same argument as above, we conclude that

$$\xi_j^* = \tilde{\xi}_j^* = \bar{\xi}_j^* + \sum'' \phi_k^*, \text{ and } \bar{\xi}_j^* = \epsilon_2 \Xi_j + \Delta_2, \qquad (\epsilon_2 = \pm 1),$$

where $\sum''$ extends over all characters ϕ_k for which $\phi_k(\tau) = -2u$, and Δ_2 is independent of j. Similarly if ξ_j' is the linear character of N such that $\bar{\xi}_j'(\tau) = 1$ and its restriction on Z_2 induces ξ_j of N_2, we have

$$\xi_j^* = \bar{\xi}_j'^* + \sum' \phi_k^*, \text{ and } \bar{\xi}_j'^* = \epsilon_3 \Xi_j + \Delta_3 \qquad (\epsilon_3 = \pm 1).$$

Let η and ϕ be characters of $H'T/T$, and assume that η is the principal character. Then we have

$$\begin{aligned}\eta^* &= \bar{\theta}^* + 3 \sum \bar{\theta}_i^* + \sum \bar{\xi}_j'^* \\ &= 1 + \delta_1 X_1 + \delta_2 X_2 + 3\epsilon_1 \sum \Theta_i + \epsilon_3 \sum \Xi_j + u(\Delta_1 + \Delta_3), \\ \phi^* &= 2 \sum \bar{\xi}_j^* = 2\epsilon_2 \sum \Xi_j + 2u\Delta_2.\end{aligned}$$

The character $\eta - \phi$ of $H'T$ induces a character ψ_0 of N. Since ψ_0 vanishes outside of $H'T$, we may apply Lemma 4. Since $\psi_0^* = \eta^* - \phi^*$, we get $w(\eta^* - \phi^*) = w(\psi_0) = 4u$. Using Lemma 4 we see that

$$\int_G (\tilde{\theta}^* - \bar{\theta}_i^*)(\bar{\Xi}_j - \bar{\Xi}_k) = \pm \int_{N_2} \psi(\bar{\xi}_j - \bar{\xi}_k) = 0.$$

Hence X_1 and X_2 are different from Ξ_j, and hence the contributions to $w(\eta^* - \phi^*)$ by the characters $1, X_1, X_2, \Theta_i, \Xi_j$ are at least $1, 1, 1, 9(u-1)/3, u$ respectively. Since $1 + 1 + 1 + 3(u-1) + u = 4u = w(\eta^* - \phi^*)$, these values are exact contributions. Hence we conclude that $\epsilon_2 = \epsilon_3$ and $2\Delta_2 = \Delta_1 + \Delta_3$. Comparing the degrees of η^* and ϕ^*, we get

$$2\epsilon_2 \sum_j \mathfrak{D}_g \Xi_j = 1 + \delta_1 \mathfrak{D}_g X_1 + \delta_2 \mathfrak{D}_g X_2 + 3\epsilon_1 \sum_i \mathfrak{D}_g \Theta_i + \epsilon_3 \sum_j \mathfrak{D}_g \Xi_j.$$

Each Ξ_j has the same degree, and $\mathcal{D}_g\Theta_i$ is also independent of i. Hence using the facts that $\epsilon_2 = \epsilon_3$ and $1 + \delta_1\mathcal{D}_gX_1 + \delta_2\mathcal{D}_gX_2 = \epsilon_1\mathcal{D}_g\Theta_i$, we have $\epsilon_2\mathcal{D}_g\Xi_j = \epsilon_1\mathcal{D}_g\Theta_i$, or $\epsilon_1 = \epsilon_2$ and $\mathcal{D}_g\Theta_i = \mathcal{D}_g\Xi_j$. If π is any character of Z_2, π^* is either θ^*, θ_i^*, or ξ_j^*. Hence the degrees of 1, X_1, X_2, Θ_i, Ξ_j are not divisible by u, but degrees of all other characters of G are multiples of u. Since $\Theta_i = \Theta_j$ on every element of $H'T$, Θ_i is not an exceptional character associated with ϕ_k. Similarly each Ξ_j is not. If X_1 is one of the exceptional characters associated with the ϕ_k of N, then only X_1 and X_2 are exceptional with respect to these ϕ_k. Hence $(v-1)/u = 2$, or $v = 2u+1$. Since $u \equiv 1 \pmod 3$, v must be divisible by 3. This is impossible, because any prime divisor of v is larger than u. Hence X_1 and X_2 are non-exceptional for ϕ_k.

Since $\xi_j^* = \bar{\xi}_j^* + \sum'' \phi_k^* = \bar{\xi}_j'^* + \sum' \phi_k^*$ we have now

$$\epsilon_2\Xi_j + \Delta_2 + \delta\sum'' E_k + \tfrac{1}{2}n\Delta = \epsilon_3\Xi_j + \Delta_3 + \delta\sum' E_k + \tfrac{1}{2}n\Delta,$$

or $\Delta_2 + \delta\sum'' E_k = \Delta_3 + \delta\sum' E_k$. Hence $\Delta_3 = \Delta_2 - \delta\sum' E_k + \delta\sum'' E_k$, and since $\Delta_1 + \Delta_3 = 2\Delta_2$,

$$\Delta_1 = \Delta_2 + \delta\sum' E_k - \delta\sum'' E_k.$$

Now

$$\begin{aligned}
\bar{\theta}^* &= 1 + \delta_1X_1 + \delta_2X_2 + \delta\sum' E_k - \delta\sum'' E_k + \Delta_2,\\
\bar{\theta}_i^* &= \epsilon\Theta_i \qquad\qquad + \delta\sum' E_k - \delta\sum'' E_k + \Delta_2,\\
\bar{\xi}_j^* &= \epsilon\Xi_j \qquad\qquad + \qquad\qquad \Delta_2,\\
\bar{\xi}_j'^* &= \epsilon\Xi_j \qquad\qquad - \delta\sum' E_k + \delta\sum'' E_k + \Delta_2,\\
\phi_k^* &= \delta E_k \qquad\qquad\qquad\qquad + \Delta\ ,
\end{aligned}$$

where we put $\epsilon_1 = \epsilon_2 = \epsilon_3 = \epsilon$.

Let τ be the generator of T, and ρ the generator of UT. We can compute the values of characters for τ and ρ. This is shown in table II.

Denote by $\langle\tau\rangle$ and $\langle\rho\rangle$ the conjugate classes of G containing τ and ρ

	τ	ρ
1	1	1
Θ_i	3ϵ	$\epsilon\theta_i(\rho)$
Ξ_j	$-\epsilon$	$\epsilon\xi_j(\rho)$
E_k	$\delta\phi_k(\tau) = \pm 2u\delta$	0
X_i	δ_i	δ_i
others	0	0

TABLE II.

respectively. Then if a and b are the coefficients of $\langle\tau\rangle$ and $\langle\rho\rangle$ resp. in the expansion of $\langle\tau\rangle^2$, we have $a=2v$ and $b=0$. By Lemma 3,

$$2v = g(4uv)^{-2}\{1+\delta_1(\mathfrak{D}_g X_1)^{-1}+\delta_2(\mathfrak{D}_g X_2)^{-1}+(8\epsilon u-9\epsilon)(\mathfrak{D}_g\Theta_i)^{-1}\}$$

and

$$0 = g(4uv)^{-2}\{1+\delta_1(\mathfrak{D}_g X_1)^{-1}+\delta_2(\mathfrak{D}_g X_2)^{-1} + (9\epsilon\sum_i\theta_i(\rho)+\epsilon\sum_j\xi_j(\rho))(\mathfrak{D}_g\Theta_i)^{-1}\}.$$

We used the equality $\mathfrak{D}_g\Theta_i=\mathfrak{D}_g\Xi_j$. But by the orthogonality relations we have

$$3+\sum_i \mathfrak{D}_g\theta_i\cdot\theta_i(\rho)=0, \quad \text{or} \quad \sum_i\theta_i(\rho)=-1,$$

and

$$3+\sum_i \mathfrak{D}_g\theta_i\cdot\theta_i(\rho)+\sum_j \mathfrak{D}_g\xi_j\cdot\xi_j(\rho)=0, \quad \text{or} \quad \sum_j\xi_j(\rho)=0.$$

Hence

$$0 = g(4uv)^{-2}\{1+\delta_1(\mathfrak{D}_g X_1)^{-1}+\delta_2(\mathfrak{D}_g X_2)^{-1}-9\epsilon(\mathfrak{D}_g\Theta_i)^{-1}\}.$$

Together with the first relation, we obtain $g=4uv^3\epsilon\,\mathfrak{D}_g\Theta_i$. Since $g=4uvm$ and $(m,uv)=1$, we conclude that $v=1$, which is clearly impossible. Hence U must be the unit group and our theorem is thus proved.

As a corollary we have

THEOREM. *Let G be a finite simple group. If the order of G is $p_1^2p_2\cdots p_n$ (p_i are primes and $p_1<p_2<\cdots<p_n$) then $G\cong LF(2,p_n)$.*

5. In this section we shall prove the following Theorem B.

THEOREM B. *Let G be a non-solvable group of even order. If we assume that all l-Sylow subgroups are cyclic when $l>2$, and the 2-Sylow subgroups are of dihedral type, then G contains a normal subgroup G_1 such that $[G:G_1]\leqq 2$ and $G_1=Z\times L$, where Z is a Z-group and L is isomorphic with some $LF(2,p)$.*

Proof. We use an induction on the order of G. Throughout the proof we may assume that the order of G is divisible by 8, since the other case is covered by Theorem A. If G contains a proper non-solvable normal subgroup U, U contains a characteristic subgroup L isomorphic with some $LF(2,p)$. L is a normal subgroup of G, so that the centralizer Z of L in G is also normal, and by Lemma 2 $Z\cup L$ has index 1 or 2. The assertion of Theorem B is then proved by putting $Z\cup L=G_1$.

We assume that all proper normal subgroups of G are solvable. From

this assumption we conclude in particular that G is perfect. By Theorems 5 and 6 of Zassenhaus [13], a solvable normal subgroup of G contains a normal subgroup P of G such that the order of P is either an odd prime or a power of 2. Since l-Sylow subgroups of G are cyclic and G is perfect, G does not contain any normal subgroup of order l by Theorem 15 of Zassenhaus [13]. Suppose now that G contains a normal subgroup P of order 2^μ. We may assume that P is a minimal one. Since 2-Sylow subgroups of G are of dihedral type, the order of P is either 2 or 4. If $[P:e]=2$, then P is in the center of G. By a theorem of Grün (see [8], or [14], p. 134), G is not perfect. (Note that S_2 is a direct factor of its normalizer, since we have assumed $[S_2:e] \geqq 8$.) If $[P:e]=4$, we consider the centralizer V of P in G. $V \neq G$, and G/V is isomorphic with a subgroup of the octahedral group, and so is solvable. In either cases we get a contradiction to our assumption of the perfectness of G. Hence G is a simple group.

Since G is perfect, elements of order 2 in G form a single conjugate class, as is seen from Grün's theorem (loc. cit.). Let τ be an element of order 2, and put $T=\{\tau\}$. Then the normalizer N of T in G contains a 2-Sylow subgroup S_2 of G and a normal subgroup H such that $HS_2=N$ and $H \cap S_2 = e$. This is proved by using Grün's theorem as before.

Consider the normalizer N_2 of S_2 in G. By our assumption $[S_2:e] \geqq 8$, and so $N_2 = S_2 \times W$. By way of contradiction, we assume that $W \neq e$. Let q be the smallest prime divisor of $[W:e]$, and W_q a q-Sylow subgroup of W. Consider the normalizer U and the centralizer V of W_q in G. Assume that $U \neq V$. Then V is a normal subgroup of U and the factor group U/V is a cyclic group of order less than q. Since $V \supseteq S_2$, we have $U = V\bar{N}_2$, where $\bar{N}_2$ is the normalizer of S_2 in U, i.e. $\bar{N}_2 = N_2 \cap U$. By the second isomorphism theorem, $[U:V] = [\bar{N}_2 : \bar{N}_2 \cap V]$, and hence $[U:V]$ is a divisor of the order of N_2. This contradicts with the choice of q. Hence $U=V$; in other words, W_q is in the center of its normalizer. Hence by a theorem of Burnside, G is not simple. We have thus proved that $W=e$, or $N_2=S_2$.

S_2 contains a cyclic subgroup Z of index 2. Since $N_2=S_2$, the normalizer D of Z in G is a dihedral group. Hence if $N=D$, our assertion follows from Lemma 7.

By way of contradiction we assume that $N \neq D$. The subgroup H is a 2-Sylow complement of N. Put $K=H \cap D$; K is cyclic. Take an element σ of order 2 in S_2, which is not contained in Z. Then K is contained in the commutator subgroup of $\{H, \sigma\}$. This implies that K is a normal subgroup of H.

The subgroup HZ is a Z-group and hence its commutator subgroup I is

cyclic. Since KZ is a maximal abelian subgroup of HZ, we have $I \cup KZ = HZ$, or $I \cup K = H$. Since we have assumed that $N \neq D$, some Sylow subgroup, say the q-Sylow subgroup I_q, of I is not contained in K. I_q is a normal subgroup of N, and Z is not contained in the centralizer of I_q, since $I_q \not\subset K = H \cap D$. But the centralizer of I_q is a normal subgroup of N with cyclic factor group, and hence it contains at least three elements of order 2. On the other hand, I_q is the totality of elements of N with orders which are powers of q. Hence the normalizer of I_q is larger than N. Let M be a maximal subgroup of G containing the normalizer of I_q in G. N must be a proper subgroup of M.

Suppose first that M is not solvable. Then we may apply our inductive hypothesis on M. M contains a normal subgroup M_1 such that $[M:M_1] \leqq 2$ and $M_1 = Z_1 \times L_1$, where Z_1 is a Z-group and $L_1 \cong LF(2,p')$ for some prime number p'. I_q is a subgroup of Z_1, since I_q commutes with at least three elements of order 2. Now if $M = M_1$, then I_q is in the centralizer of S_2 in G. This is impossible. Hence $[M:M_1] = 2$. Since M/Z_1 is a linear group by Lemma 2, we have $[N:Z_1] = 2(p' - \epsilon)$, where $\epsilon = (-1)^{(p'-1)/2}$. By the inductive hypothesis, the order of Z_1 is relatively prime to $p'^2 - 1$, a divisor of the order of L_1, and hence the order of N can not be divisible by $(p' + \epsilon)/2$. This gives a contradiction, since M contains an element π of order 2 such that $\pi \notin L_1$ and π commutes with an element of order $(p' + \epsilon)/2$. Hence M must be solvable.

Now by Theorem 6 of Zassenhaus [13], M contains a normal subgroup M_1 with the properties that M_1 is a Z-group and M/M_1 either has order 2 or is isomorphic with the tetrahedral or octahedral group.

If M/M_1 is of order 2, then M contains a 2-Sylow complement J. J is a normal subgroup of M and contains H as a proper subgroup. As before, K is a normal subgroup of J, since K is in the commutator subgroup of a Z-group $\{J, \sigma\}$. Similarly, I is normal in J, since I is in the commutator subgroup of JZ. Hence $H = I \cup K$ is normal in J. Since J is a Z-group, every normal subgroup of J is characteristic. Hence H is a normal subgroup of M. Consider the factor group M/H and the image $\bar{\tau}$ of the generator τ of the subgroup T of order 2 under the natural homomorphism. The normalizer of $\bar{\tau}$ in M/H is N/H, and since $H = J \cap N$, $\bar{\tau}$ induces the automorphism of J/H which maps every element into its inverse (see Lemma 1 of Zassenhaus [12]). Now N/H is isomorphic with S_2 and is generated by two elements ρ and ρ' of order 2. If we take a Sylow subgroup $\bar{P}$ of J/H, ρ (or $\rho\bar{\tau}$) commutes with all elements of P. Similarly ρ' (or $\rho'\bar{\tau}$) commutes with all elements of P. Hence $\bar{P} \cup N/H = \bar{P} \times N/H$, and this is a contradiction.

The 2-Sylow subgroup S_2 coincides with its own normalizer. Hence M can not contain any proper normal subgroup with odd index. In particular M/M_1 can not be isomorphic with the tetrahedral group.

Suppose finally that M/M_1 is isomorphic with the octahedral group. Then M contains a maximal subgroup M_2 such that $[M:M_2]=2$ and M_2/M_1 is isomorphic with the tetrahedral group. A 2-Sylow subgroup of M_2 is a dihedral group and its three maximal subgroups are conjugate in M_2 to each other. Hence the order of M_2 is not divisible by 8, or M_1 is of odd order. Consider a Sylow subgroup S of M_1, and denote by X and Y the normalizer and the centralizer of S in M_2 respectively. Y is a normal subgroup of X and the factor group X/Y is cyclic, since S is cyclic. Put $Q=M_2\cap S_2$. Then Q is a non-cyclic group of order 4. Now $M_2=M_1X$, and so $X/X\cap M_1\cong M_2/M_1$. Hence $Y\cup M_1$ is a normal subgroup of M_2 with cyclic factor group $M_2/Y\cup M_1$. Since M_2/M_1 is isomorphic with the tetrahedral group, the order of $Y\cup M_1$ is divisible by 4, and so $Y\cup M_1\supseteq Q=M_2\cap S_2$. This means that Q is contained in the centralizer of some conjugate subgroup of S. Since S was any Sylow subgroup of M_1, Q is in the centralizer of M_1, or $Q\cup M_1=Q\times M_1$, and Q is a normal subgroup of M. The factor group M/Q is of dihedral type, because S_2 is its own normalizer in G. This implies in particular that M_2/Q is cyclic. M_2 contains a subgroup R such that $RQ=M_2$ and $R\cap Q=e$. Then R is isomorphic with M_2/Q, and $R\supseteq M_1$. We want to prove that the normalizer, in G, of any subgroup of R is contained in M. Consider a subgroup X of R. If $X\cap M_1\neq e$, the normalizer of X is contained in the normalizer of $X\cap M_1$. Since M is a maximal subgroup of G and G is simple, the normalizer of $X\cap M_1$ is M, and so the normalizer of X is contained in M. If $X\cap M_1=e$, then the centralizer of X is of odd order. Hence the normalizer Y of X is a Z-group. Since R is contained in the commutator subgroup of Y, R is a normal subgroup of Y, or Y is contained in the normalizer of M_1, i.e. $Y\subseteq M$. Hence all conjugate classes of M containing elements $\neq 1$ are special with respect to G.

Consider the irreducible characters of M and M_1Q. Set $[M_1:e]=v$. Let θ_0 be the principal character of M_1Q, ω a non-principal character of M_1Q/M_1, and $\theta_1,\cdots,\theta_q$ non-principal characters of M_1Q/Q, one taken from each associated class of characters, where $q=(v-1)/2$.

If ψ is any character of M_1Q, we denote by $\bar{\psi}$ that character of M which is induced by ψ. The following relations are obtained:

$$\bar{\theta}_0=1+\eta+2\phi_0, \qquad \bar{\omega}=\xi_1+\xi_2,$$

$$\bar{\theta}_i=\phi_{3i-2}+\phi_{3i-1}+\phi_{3i}, \qquad \overline{\omega\theta_i}=\Theta_i,$$

where 1 is the principal character of M, η the non-principal linear character, $\phi_0, \cdots, \phi_{3q}$ characters of degree 2, ξ_1 and ξ_2 characters of M/M_1 of degree 3, and the Θ_i are of degree 6. All these characters of M are irreducible.

We apply Lemma 5. G has exceptional characters $\Phi_0, \cdots, \Phi_{3q}$ associated with $\phi_0, \cdots, \phi_{3q}$. If η_0 is the principal character of M_2, then $\bar{\eta}_0 = 1 + \eta$. Hence $\int_G |\eta_0^* - \phi_i^*|^2 = 3$. Since $\phi_i^* = \epsilon\Phi_i + \Delta_1$, $\eta_0^* = 1 + \cdots$ and $3q \geqq 3$,

$$\phi_i^* = \epsilon\Phi_i + a\sum\Phi_k + aX + \Delta_2 \qquad (\epsilon = \pm 1),$$

and

$$\eta_0^* = 1 + a\sum\Phi_k + (a+\epsilon)X + \Delta_2,$$

where Δ_2 does not contain X or Φ_k. (The multiplicity of X in ϕ_i^* is the same as that of Φ_k in η_0^*, since $w(\phi_i^*) = w(\eta_0^*) - 1$.) Hence

$$\theta_0^* = 1 + 2\epsilon\Phi_0 + \epsilon X + 3a\sum\Phi_k + 3aX + 3\Delta_2,$$

$$\theta_i^* = \epsilon(\Phi_{3i-2} + \Phi_{3i-1} + \Phi_{3i}) + 3a\sum\Phi_k + 3aX + 3\Delta_2.$$

Applying Lemma 4, we have

$$w((\omega\theta_i)^* - \theta_0^*) = w(\Theta_i - \bar{\theta}_0) = 7,$$

$$w((\omega\theta_i)^* - \theta_j^*) = w(\Theta_i - \bar{\theta}_j) = 4.$$

Moreover, by making computations similar to those used in the proof of Lemma 4, we obtain

$$\int_G (\omega\theta_i)^*(\bar{\Phi}_j - \bar{\Phi}_k) = \pm\int_M \Theta_i(\bar{\phi}_j - \bar{\phi}_k) = 0.$$

Suppose $(\omega\theta_i)^*$ contains Φ_k and X with the multiplicities b and c respectively. Then

$$7 = 1 + (3a-b)^2 3q + (3a+2\epsilon-b)^2 + (3a+\epsilon-c)^2 + \gamma,$$

$$4 = (3a-b)^2(3q-2) + 3(3a+\epsilon-b)^2 + (3a-c)^2 + \gamma,$$

where γ is a non-negative integer. Hence

$$6 = 2(3a-b)^2 + (3a+2\epsilon-b)^2 + (3a+\epsilon-c)^2 + \gamma',$$

$$4 = 3(3a+\epsilon-b)^2 + (3a-c)^2 + \gamma',$$

where $\gamma' = \gamma + (3a-b)^2(3q-2)$. We have two cases:

Case I. $3a - b = 0$. In this case, the above equation can be read as

$$2 = (3a+\epsilon-c)^2 + \gamma' \text{ and } 1 = (3a-c)^2 + \gamma'.$$

Hence $3a-c=0$ and $\gamma'=1=\gamma$.

Case II. $3a-b+\epsilon=0$. In this case,

$$3=(3a+\epsilon-c)^2+\gamma' \text{ and } 4=(3a-c)^2+\gamma'.$$

Hence

$$3=(3a-c)^2+2\epsilon(3a-c)+1+\gamma'=5+2\epsilon(3a-c),$$

or $3a-c=-\epsilon$ and $\gamma'=3$. Since $\gamma'=\gamma+3q-2$, we conclude that $v=3$ and $\gamma=2$.

We shall treat these cases separately.

Case I. If $v>3$, G has exceptional characters $E_1,\cdots,E_q$ associated with $\Theta_1,\cdots,\Theta_q$ (see Lemma 5). If $v=3$ we cannot apply Lemma 5; but there exists only one character E_1 of G whose multiplicity in $(\omega\theta_1)^*$ is not divisible by 3, since $\gamma=1$. We take E_1 as the exceptional character associated with Θ_1. In any case we have decompositions of induced characters in the following form:

$$\theta_0^*=1+2\epsilon\Phi_0+\epsilon X+3a\sum\Phi_k+3aX+3b\sum E_l+3\Delta_3,$$

$$\theta_i^*=\epsilon(\Phi_{3i-2}+\Phi_{3i-1}+\Phi_{3i})+3a\sum\Phi_k+3aX+3b\sum E_l+3\Delta_3 \quad (i\geqq 1),$$

$$(\omega\theta_j)^*=\epsilon E_j+3a\sum\Phi_k+3aX+3b\sum E_l+3\Delta_3 \quad (j\geqq 1).$$

Consider then ω^*; we have from Lemma 4

$$w(\omega^*-\theta_0^*)=8,\ w(\omega^*-\theta_i^*)=5 \text{ and } w(\omega^*-(\omega\theta_j)^*)=3.$$

Since $\int_G\omega^*(\Phi_i-\Phi_j)=\pm\int_M\overline{\omega}(\phi_i-\phi_j)=0$ and since

$$\int_G\omega^*(E_i-E_j)=\pm\int_M\overline{\omega}(\Theta_i-\Theta_j)=0,$$

we have the decomposition of ω^* as follows:

$$\omega^*=a'\sum\Phi_k+a''X+b'\sum E_l+\Delta_4,$$

where Δ_4 does not contain Φ_k, X and E_l.

We now get three equations for a', a'', and b', namely

$$3=(a'-3a)^2(3q+1)+(a''-3a)^2+(b'-3b)^2(q-1)+(b'-3b-\epsilon)^2+\Gamma,$$

$$5=(a'-3a)^2(3q-2)+3(a'-3a-\epsilon)^2+(a''-3a)^2+(b'-3b)^2q+\Gamma,$$

$$8=1+(a'-3a)^23q+(a'-3a-2\epsilon)^2+(a''-3a-\epsilon)^2+(b'-3b)^2+\Gamma,$$

where $\Gamma = w(\Delta_4 - 3\Delta_3) \geqq 0$. The first two equations may be written as

$$3 = 3(a' - 3a)^2 + (b' - 3b - \epsilon)^2 + \Gamma',$$

$$5 = 3(a' - 3a - \epsilon)^2 + (b' - 3b)^2 + \Gamma' \qquad (\Gamma' \geqq 0).$$

Hence $a' = 3a$ or $a' = 3a + \epsilon$. But if $a' = 3a + \epsilon$, then the first equation shows that $b' = 3b + \epsilon$ and $\Gamma' = 0$. This solution does not satisfy the second one. Hence we have $a' = 3a$ and $b' = 3b$. If we use these values for a' and b', then the first and the last equations become

$$3 = (a'' - 3a)^2 + 1 + \Gamma, \qquad 8 = 1 + 4 + (a'' - 3a - \epsilon)^2 + \Gamma.$$

Hence $2 = (a'' - 3a)^2 - 2\epsilon(a'' - 3a) + \Gamma = 2 - 2\epsilon(a'' - 3a)$, or $a'' = 3a$, and hence $\Gamma = 2$. Now we can write

$$\omega^* = 3a \sum \Phi_k + 3aX + 3b \sum E_l + \Delta_4, \qquad \Delta_4 = \delta_1 Y + \delta_2 Z + 3\Delta_3,$$

where $\delta_i = \pm 1$ and Y and Z are irreducible characters of G, which are not Φ_k, X, E_l $(k = 0, \cdots, 3q;\ l = 1, 2, \cdots, q)$.

We can now compute the values of the characters. Let τ be an element of order 2, and ρ a generator of the subgroup R.

	1	τ	ρ
1	1	1	1
E_j	$6\epsilon + 12bv$	-2ϵ	0
Φ_i	$2\epsilon + 12av$	2ϵ	$\epsilon\phi_i$
X	$\epsilon + 12av$	ϵ	ϵ
Y	$3\delta_1 + 12yv$	$-\delta_1$	0
Z	$3\delta_2 + 12zv$	$-\delta_2$	0

TABLE III.

All other irreducible characters of G, not listed in Table III, vanish on ρ. We compute $\langle\tau\rangle^2$ in the group ring; then $\langle\rho\rangle$ has multiplicity $3v$ in $\langle\tau\rangle^2$, since the normalizer of R contains exactly $3v$ elements of order 2. Hence by Lemma 3 we have

$$3v = g(8v)^{-2}\{1 + \epsilon(\mathfrak{D}_g X)^{-1} + 4\epsilon(\mathfrak{D}_g \Phi_i)^{-1} \sum_i \phi_i(\rho)\}.$$

If we put $x = \mathfrak{D}_g X$, then $\mathfrak{D}_g \Phi_i = x + \epsilon$. An orthogonality relation yields $\sum_i \phi_i(\rho) = -1$. Hence

$$3v = g(8v)^{-2}\{1 + \epsilon x^{-1} - 4\epsilon(x+\epsilon)^{-1}\} = g(8v)^{-2}(x-\epsilon)^2 x^{-1}(x+\epsilon)^{-1},$$

or $g = 3 \cdot 8^2 v^3 (12av + \epsilon)(12av + 2\epsilon)(12av)^{-2}$, that is,

$$g = 24v \cdot (12av + \epsilon)(6av + \epsilon)/3^2 a^2.$$

This is impossible, since the last fraction should be an integer.

Case II. In this case we have $v = 3$, and the induced characters are decomposed as follows:

$$\theta_0^* = 1 + 2\epsilon\Phi_0 + \epsilon X + 3\Delta_1, \qquad \theta_1^* = \epsilon(\Phi_1 + \Phi_2 + \Phi_3) + 3\Delta_1,$$

$$(\omega\theta_1)^* = \epsilon(\Phi_0 + \Phi_1 + \Phi_2 + \Phi_3) + \epsilon X + \delta_1 Y + \delta_2 Z + 3\Delta_1,$$

where $\delta_i = \pm 1$ and Y, Z are irreducible characters of G, not equal to X, Φ_i, or the principal character. In the same way as in Case I, the induced character ω^* is decomposed into a sum of irreducible characters as follows: $\omega^* = a' \sum \Phi_k + a'' X + \Delta_5$, where Δ_5 does not contain Φ_k or X. If we write $\Delta_1 = a \sum \Phi_k + aX + \Delta_2$ as before, we get three equations for a' and a''. The equation $w(\omega^* - (\omega\theta_1)^*) = 3$ implies that

$$3 = 4(a' - 3a - \epsilon)^2 + a'' - 3a - \epsilon)^2 + \cdots.$$

Hence a' must be equal to $3a + \epsilon$. Similarly $w(\omega^* - \theta_0^*) = 8$ shows that

$$8 = 1 + 4 + (a'' - 3a - \epsilon)^2 + w(\Delta_5 - 3\Delta_2),$$

or $(a'' - 3a - \epsilon) + w(\Delta_5 - 3\Delta_2) = 3$. Finally $w(\omega^* - \theta_1^*) = 5$ implies that

$$5 = 1 + (a'' - 3a)^2 + w(\Delta_5 - 3\Delta_2),$$

or $(a'' - 3a)^2 + w(\Delta_5 - 3\Delta_2) = 4$. Hence $a'' = 3a + \epsilon$ and $w(\Delta_5 - 3\Delta_2) = 3$.

Let b, b' be the multiplicities of Y in Δ_2 and Δ_5 respectively, and similarly let c and c' be those of Z. Then $w(\Delta_5 - 3\Delta_2) = 3$ implies that

$$(3b - b')^2 + (3c - c')^2 + \Gamma = 3 \text{ or, since } w(\omega^* - (\omega\theta_1)^*) = 3,$$

we get $(3b + \delta_1 - b')^2 + (3c + \delta_2 + c')^2 + \Gamma = 3$. Hence $3b = b'$ or $3c = c'$. Without loss of generality we may assume that $b' = 3b$. Then $c' = 3c + \delta_2$ and $\Gamma = 2$. Hence

$$\omega^* = \epsilon \sum \Phi_k + \epsilon X + \delta_2 Z + \delta_3 V + \delta_4 W + 3\Delta_1,$$

where $\delta_3, \delta_4 = \pm 1$ and V, W are irreducible characters of G. We have the following table. (τ and ρ have the same meanings as in Table III.)

	1	τ	ρ
1	1	1	1
Φ_i	$36a+11\epsilon$	$-\epsilon$	$\epsilon\phi_i(\rho)$
Y	$36y+6\delta_1$	$-2\delta_1$	0
X	$36a+10\epsilon$	-2ϵ	ϵ
Z	$36z+9\delta_2$	$-3\delta_2$	0
V	$36v+3\delta_3$	$-\delta_3$	0
W	$36w+3\delta_4$	$-\delta_4$	0

TABLE IV.

All characters of G which do not appear in this table have degrees divisible by 36, and hence vanish on τ and ρ. From this table we get, by Lemma 3, the following equation:

$$9 = g(24)^{-2}\{1+4\epsilon(\mathcal{D}_g X)^{-1}+\epsilon(\mathcal{D}_g\Phi_i)^{-1}\sum_i \phi_i(\rho)\}.$$

Hence $9 = g(24)^{-2}\{1+4\epsilon(36a+10\epsilon)^{-1}-\epsilon(36a+11\epsilon)^{-1}\}$, since $\sum_i \phi_i(\rho) = -1$; so that

$$g = 9\cdot 8\cdot(36a+11\epsilon)(18a+5\epsilon)(3a+\epsilon)^{-2}.$$

It can be seen easily that $3a+\epsilon$ is relatively prime to $36a+11\epsilon$ and $18a+5\epsilon$. Hence $(3a+\epsilon)^2$ must be 1, i.e. $3a+\epsilon = \pm 1$, or $a=0$. Hence $\epsilon = +1$ and $g = 8\cdot 9\cdot 5\cdot 11 = 3960$.

We shall complete the first column of the table IV. Since $\theta_1{}^*$ and $(\omega\theta_1)^*$ have the same degree, we get

$$\epsilon\mathcal{D}_g\Phi_i+\epsilon\mathcal{D}_g X+\delta_1\mathcal{D}_g Y+\delta_2\mathcal{D}_g Z = 0,$$

or $\delta_1 y+\delta_2 z = -1$. Similarly, since $\mathcal{D}_g(\omega\theta_4)^* = \mathcal{D}_g\omega^*$,

$$\delta_1\mathcal{D}_g Y = \delta_3\mathcal{D}_g V+\delta_4\mathcal{D}_g W,$$

or $\delta_1 y = \delta_3 v+\delta_4 w$. If one of y, z, v, w is greater than 1, then G has a representation of degree at least equal to $72-9 = 63$. Since $(63)^2 = 3969 > 3960 = g$, this can not happen. Hence each of the integers y, z, v, w is either 0 or 1. If $z=1$, then y must be 0 and $\delta_2 = -1$. Hence $\mathcal{D}_g Z = 27 = 3^3$, which is impossible. (For the order of G is not divisible by 27.) Hence $z=0$. So we conclude that $\delta_2 = 1$, $y=1$, and $\delta_1 = -1$, i.e., $\mathcal{D}_g Y = 30$ and $\mathcal{D}_g Z = 9$. Since $\delta_3 v+\delta_4 w = -1$, we may assume with-

out loss of gnerality that $v=0$, $\delta_3=1$, $w=1$, and $\delta_4=-1$. Hence $\mathfrak{D}_g V=3$ and $\mathfrak{D}_g W=33$. The sum of the squares of the degrees of all the irreducible characters is equal to the order of G, and the degrees of the missing characters are multiples of 36. Hence

$$g=1+4(11)^2+(10)^2+(30)^2+(9)^2+(3)^2+(33)^2+k(36)^2,$$

or $k=1$. Hence there is only one character missing and the degree of the missing character is exactly 36. This is however impossible, since characters of degree 36 belong to a 2-block of defect 1, which contains exactly two characters of degree 36 (see [1]).

In both cases we have derived a contradiction, and hence the normalizer of elements of order 2 must be of dihedral type. This proves Theorem B.

6. We shall now prove the following theorem:

THEOREM C. *Let G be a non-solvable group of finite even order. We assume that all Sylow subgroups of odd order are cyclic and that a 2-Sylow subgroup is a generalized quaternion group. Then G contains a normal subgroup G_1 with the properties that $[G:G_1]\leqq 2$ and G_1 is a direct product of a Z-group and a subgroup isomorphic with $SL(2,p)$ for some prime number p.*

Proof. We shall again use the inductive argument. Assume that G is not perfect. Then the commutator subgroup G' of G is a proper subgroup satisfying the assumptions made on G. Hence we may apply the inductive hypothesis to G'. G' has the structure described in Theorem C; in particular, G' contains a characteristic subgroup T of order 2. T is clearly a normal subgroup of G. Now the factor group G/T satisfies all the assumptions of Theorem B. Hence G/T contains a normal subgroup G_1/T such that $[G:G_1]\leqq 2$ and $G_1/T\cong \bar{Z}\times \bar{L}$, where $\bar{Z}$ is a Z-group of odd order and $\bar{L}$ is isomorphic with $LF(2,p)$. Hence G_1 contains two normal subgroups L and Z such that $G_1=IZ$, $L\cap Z=T$, $\bar{L}\cong L/T$, and $\bar{Z}\cong Z/T$. Since T is normal and the 2-Sylow subgroups are of quaternion type, L contains only one subgroup of order 2. Hence by Lemma 1, L is isomorphic with $SL(2,p)$. Since $\bar{Z}$ is of odd order, Z contains a subgroup Z_1 such that $Z=Z_1\times T$. Z_1 is a normal subgroup of G_1 and $G_1=Z_1\times L$. This proves our theorem in this case.

We may, therefore, assume that G is perfect. Take a 2-Sylow subgroup S_2 of G. The center T of S_2 is of order 2. We denote by τ the generator of T and consider the normalizer N of τ in G. Since G is 2-normal in the sense of Grün [8], we have an isomorphism $G/G'(2)\cong N/N'(2)$, where

$G'(2)$ and $N'(2)$ are the 2-commutator subgroups of G and N respectively. Since we have assumed that G is perfect, N does not contain any normal subgroup of index 2.

Suppose that N is solvable. Using a theorem of Zassenhaus, we know that N contains a normal subgroup H such that N/H is isomorphic with the tetrahedral group. Hence $H \cap S_2 = T$. Now H is a Z-group, so that H contains a subgroup I such that $H = I \times T$. Let J be the union of I and S_2. Consider a Sylow subgroup X of I, the normalizer Y of X in N, and the centralizer Z of X. Then we have $N = IY$ and Z is a normal subgroup of Y with cyclic factor group Y/Z. Hence ZI is a normal subgroup of N with a cyclic factor group. This implies that the index $[N:ZI]$ is odd, and hence S_2 is in the centralizer of some conjugate subgroup of X. Since X was any Sylow subgroup of I, we conclude that $J = I \cup S_2 = I \times S_2$. Hence J must be the centralizer of S_2 in G. Now N contains a subgroup K such that $KS_2 = N$, $K \cap J = I$, and $[K:I] = 3$. If the order of I were divisible by 3, then the normalizer and the centralizer of a 3-Sylow subgroup would coincide with each other, and so G would not be perfect. Hence the order of I is not a multiple of 3. Consider then a subgroup P of K with order 3. P is contained in a 3-Sylow subgroup S_3 of G. Let N_3 and Z_3 be the normalizer and the centralizer of S_3 in G respectively. Since G is perfect, Z_3 is a normal subgroup of N_3 with index 2. If $[Z_3:e]$ were odd, there would be an element σ of order 2 such that σ induces an automorphism of S_3 which maps every element into its inverse. Hence the normalizer of P in G would contain two elements of order 2, σ and τ, such that one of them commutes with every element in P but the other does not. This is impossible, since σ and τ are conjugate in the normalizer of P. Hence the order of Z_3 is even. This implies that G contains an element ν of order 4 such that $\nu\pi\nu^{-1} = \pi^{-1}$ for every $\pi \in P$. The normalizer of P in G contains ν, and ν^2 is conjugate to τ in this normalizer. This is clearly not the case, because of the structure of N. Hence N is not solvable.

The factor group N/T satisfies all the assumptions of Theorem B, but does not contain any normal subgroup of index 2. Hence as before N is a direct product of a Z-group Z and a linear group. If $Z \neq e$, we consider a Sylow subgroup X of Z which belongs to the smallest prime divisor of the order of Z. The normalizer Y of X in G is a non-solvable proper subgroup of G. Hence we may apply the inductive hypothesis and conclude that T is in the center of Y; in other words, Y is contained in N. From the choice of X, we see that Y coincides with the centralizer of X. Hence by a theorem of Burnside, G is not perfect. Hence we have $Z = e$, or N is isomorphic with

$SL(2,p)$ for some p. Now our Theorem C is a consequence of the following general proposition.

THEOREM D. *Let G be a group of finite even order. If G contains an element τ of order 2 such that the normalizer of τ in G is isomorphic with the special linear group $SL(2,p)$ for some $p>2$, then G contains an abelian normal subgroup G_0 such that $G/G_0 \cong SL(2,p)$, and if $G_0 \neq e$, then every primary component of G_0 is not cyclic.*

Proof. Denote by N the normalizer of the element τ of order 2. By assumption N is isomorphic with $SL(2,p)$. In the following we consider only the case $p \equiv 1 \pmod 4$. The other case will be treated in exactly the same way, with a slight modification on signs.

Put $p-1=2^a h$, $(h,2)=1$. Then $a \geqq 2$ and $(p+1)/2 = 2a+1$ is odd. We shall consider the structure of 2-blocks of G with positive defect. By Theorem 4 of Brauer [2], G has exactly a 2-blocks of defect 1, $(h-1)/2$ blocks of defect a, and one block of defect $a+1$. We have a one-to-one correspondence between blocks of positive defect of G and those of N. Namely if we take a block B of positive defect d $(d>0)$ of G, then B determines a block B of defect d in the normalizer $N(D)$ of the defect group D of B. Since $d>0$, D is not the unit group, and we may take D as a subgroup of N. Then T, the subgroup generated by τ, is a characteristic subgroup of D. Hence the normalizer $N(D)$ is also a subgroup of N. Hence $\bar{B}$ determines a block B^* of N. In this manner we get a one-to-one correspondence $B \leftrightarrow B^*$. We want to show that B and B^* consist of exactly the same number of irreducible characters. Consider any subgroup X of N, whose order is a power of 2, and a block $\bar{B}$ of the normalizer Y of X in G. Then $\bar{B}$ determines a block B of G and a block B' of N. Let $\bar{\omega}$ be the linear character of the center of the modular group ring of Y, corresponding to the block $\bar{B}$. Then the corresponding linear character ω for B satisfies

$$\omega(C) \equiv \sum \bar{\omega}(C_\lambda) \pmod{\mathfrak{p}}$$

where $\mathfrak{p}$ is a prime divisor of 2 in a suitable algebraic number field and C_λ ranges over all classes of Y which lie in C and whose elements belong to the centralizer of Y. Similarly the linear character ω' of B' is defined by

$$\omega'(C') \equiv \sum \bar{\omega}(C'_\mu) \pmod{\mathfrak{p}}\,; \text{ hence } \omega(C) \equiv \sum \omega'(C'_\nu) \pmod{\mathfrak{p}},$$

where C'_ν ranges over all classes of N which belong to C. This implies that $B'=B^*$ in the previous notation. Hence if a block $\bar{B}$ of X determines a block B of G, then $\bar{B}$ determines a block B^* of N. We apply Theorems 2

and 3 of Brauer [3] to G and N, and conclude that the number of characters in B is equal to that in B^*.

Let B_1 be the first 2-block of N. Then $\bar{B}_1$ consists of the principal character 1, one character ϕ_1 of degree p, two conjugate characters ϕ_2 and ϕ_3 of degree $(p+1)/2$, two characters ϕ_4 and ϕ_5 of degree $(p-1)/2$, and $2^{a-1}-1$ characters ϕ_i $(i=6,7,\cdots,2^a+4)$ of degree $p+1$ (see Section 1 of this paper). Let ϕ be a linear combination of these characters which vanishes on all 2-regular elements, say $\phi=1+\phi_1-\phi_2-\phi_3$. Since all 2-singular classes of N are special with respect to G, we may apply Lemma 4. It follows that $w(\phi^*)=w(\phi)=4$. On the other hand, ϕ^* contains the principal character of G with multiplicity 1. Hence

$$\phi^*=1+\epsilon_1X_1+\epsilon_2X_2+\epsilon_3X_3 \qquad (\epsilon_i=\pm 1).$$

Similarly we get

$$(1-\phi_2+\phi_4)^*=1+\delta_1Y_1+\delta_2Y_2 \qquad (\delta_i=\pm 1).$$

But $(1+\epsilon_1X_1+\epsilon_2X_2+\epsilon_3X_3)-(1+\delta_1Y_1+\delta_2Y_2)=(\phi_1-\phi_3+\phi_4)^*$ is a sum of three irreducible characters. Hence

$$(1-\phi_2+\phi_4)^*=1+\epsilon_2X_2+\epsilon_4X_4 \qquad (\epsilon_4=\pm 1).$$

Similarly

$$(1-\phi_3+\phi_5)^*=1+\epsilon_3X_3+\epsilon_5X_5 \qquad (\epsilon_5=\pm 1),$$

and

$$(\phi_i-\phi_2-\phi_3)^*=\epsilon_iX_i+\epsilon_2X_2+\epsilon_3X_3 \qquad (\epsilon_i=\pm 1).$$

These characters X_i are all different. Any one of the above relations between the X_i vanishes on all 2-regular classes. Hence by Theorem 6 of Brauer-Nesbitt [4], these relations remain true if we omit all characters which are not in the fixed 2-block B of G. Hence if some X_k is not in the first 2-block B_1 of G, then X_k must vanish on all 2-regular classes. This is clearly impossible. Hence all the X_k are in B_1 and B_1 contains no other characters. Using Lemma 4, we obtain the following table (for notations see Table I):

	2-regular	order 2	order 2^a	$p+1$
$1+\epsilon_1X_1+\epsilon_2X_2+\epsilon_3X_3$	0	0	4	0
$1+\epsilon_2X_2+\epsilon_4X_4$	0	$1-p$	2	2
$1+\epsilon_3X_3+\epsilon_5X_5$	0	$1-p$	2	2
$\epsilon_iX_i+\epsilon_2X_2+\epsilon_3X_3$	0	$\begin{cases}0\\-2(p+1)\end{cases}$	$2+\theta_i$	0

TABLE V.

These relations are linear equations of the values of X_k for elements of order 2, etc. We want to solve these equations. Before doing so, we shall derive other relations.

If $\bar{B}_\kappa$ is a block of defect α of N, $\bar{B}_\kappa$ consists of 2^α characters $\theta_1, \cdots, \theta_{2^\alpha}$ of degree $p+1$. They take the same value on every 2-regular class. We apply Lemma 5. The exceptional characters $\Theta_1, \cdots, \Theta_{2^\alpha}$ associated with these $\theta_1, \cdots, \theta_{2^\alpha}$ belong to the same block B_κ of G, since they are identical on 2-regular classes. This block has defect greater than 2, since $\alpha \geqq 2$. In the first 2-block B_1, we have several degree relations, for instance $1 + \epsilon_1 \mathcal{D}_g X_1 - \epsilon_i \mathcal{D}_g X_i = 0$. Hence at least one of X_1 and X_i is not exceptional. Hence B_1 contains at least 4 characters which are not exceptional for θ_i. Hence B_κ is a block of defect α. Since each block of defect α contains exactly 2^α characters, B_κ contains only exceptional characters $\Theta_1, \cdots, \Theta_{2^\alpha}$. By definition of exceptional characters, Θ_i is decomposed as $\Theta_i = \epsilon\theta_i + \Delta$ $(i = 1, 2, \cdots, 2^\alpha)$ on N. Hence if σ is any 2-singular element of G, we have $\Theta_i(\sigma) = \epsilon\theta_i(\sigma) + c$, where c is independent of i. The orthogonality relation yields

$$\sum_i (\mathcal{D}_g \Theta_i)\Theta_i(\sigma) = 0, \text{ or } \sum \Theta_i(\sigma) = 0.$$

Similarly $\sum_i \theta_i(\sigma) = 0$. Hence we have $c 2^\alpha = 0$, $c = 0$. Hence we conclude that $\Theta_i(\sigma) = \epsilon\theta_i(\sigma)$ for any 2-singular element. If we take another block $\bar{B}_\lambda$ of defect α, then the exceptional characters of G associated with characters in $\bar{B}_\lambda$ form a block B_λ of defect α, and $B_\lambda \neq B_\kappa$. Hence every block of defect α of G is obtained in the above process of taking exceptional characters.

If B is a block of defect 1 of G, B consists of two characters Ψ_1 and Ψ_2 which take the same value on every 2-regular element. Hence we have

$$2 = w(\Psi_1 - \Psi_2) = \int_N |\Psi_1 - \Psi_2|^2,$$

or $\Psi_1 - \Psi_2 = \psi_1 - \psi_2$ on N, where ψ_1 and ψ_2 are irreducible characters of N. ψ_i is not in any block of defect α. For if ψ_1 belongs to a block $\bar{B}$ of defect α, then ψ_2 will belong to $\bar{B}$. Since $(\psi_1 - \psi_2)^* = \Psi_1 - \Psi_2$, Ψ_i must be exceptional characters associated with ψ_i. But ψ_i belongs to $\bar{B}$, and so Ψ_i belongs to a block B of defect α. This is not the case. Hence ψ_i does not belong to any block of defect α. Since

$$\int_G (\Psi_1 - \Psi_2)(1 + \epsilon_2 X_2 + \epsilon_4 X_4) = \int_N (\psi_1 - \psi_2)(1 - \phi_2 + \phi_4) = 0,$$

ψ_i is not ϕ_2, and similarly ψ_i is not equal to any $\phi_i \varepsilon \bar{B}_1$. Hence ψ_1 and ψ_2 are in a block of defect 1. Hence for any 2-singular element of G, we have $\Psi_i(\sigma) = \psi_i(\sigma)$ $(i = 1, 2)$.

Let σ be any 2-singular element of G. The orthogonality relation yields $\sum_{\mu} |\chi_\mu(\sigma)|^2 = n(\sigma)$, where χ_μ ranges over all irreducible characters of G and $n(\sigma)$ is the order of the normalizer of σ. Since the normalizer of σ is contained in N, we have the same relation for N, namely, $\sum_{\nu} |\phi_\nu(\sigma)|^2 = n(\sigma)$, where ϕ_ν ranges over all irreducible charatcers of N. Now any character of defect 0 vanishes on σ. If B is not the first 2-block of G, then the contribution of the characters in B to $\sum_{\mu}$ is equal to that of the corresponding block of N to $\sum_{\nu}$. Hence we conclude that

$$(*) \qquad \sum_i |X_i(\sigma)|^2 = \sum_i |\phi_i(\sigma)|^2$$

where the summation is extended over $i = 1, 2, \cdots, 2^{a-1}+5$. We apply this relation for $\sigma = \tau, \rho, v$, where τ is an element of order 2, ρ an element of order 2^a and v of order $p+1$.

Put $x_i = \epsilon_i X_i(\tau)$. Then we have from table V

$$\begin{aligned} 1 + x_1 + x_2 + x_3 &= 0 \\ 1 + x_2 + x_4 &= 1 - p \\ 1 + x_3 + x_5 &= 1 - p \\ x_4 + x_5 + x_6 &= 0 \qquad \text{or } -2(p+1). \end{aligned}$$

Put $x_2 = s$ and $x_3 = t$. Then s and t are real. The relation (*) is now written as

$$\begin{aligned} 1 + (1+s+t)^2 + s^2 + t^2 + (p+s)^2 + (p+t)^2 + (s+t)^2(2^{a-2}-1) \\ + (s+t+2(p+1)^2(2^{a-2}) \\ = 1 + p^2 + \tfrac{1}{2}(p-1)^2 + \tfrac{1}{2}(p+1)^2 + (p+1)^2(2^{a-1}-1), \end{aligned}$$

or $(2^{a-1}+1)(p+1+s+t)^2 + (s-t)^2 = 0$. Since s and t are real, we have $s = t = -\frac{1}{2}(p+1)$. Now we obtain the first column of the following table VI.

	τ	ρ	v
1	1	1	1
x_1	$\epsilon_1 p$	ϵ_1	$-\epsilon_1$
x_2	$-\epsilon_2 \frac{1}{2}(p+1)$	ϵ_2	0
x_3	$-\epsilon_2 \frac{1}{2}(p+1)$	ϵ_3	0
x_4	$-\epsilon_4 \frac{1}{2}(p-1)$	0	ϵ_4
x_5	$-\epsilon_5 \frac{1}{2}(p-1)$	0	ϵ_5
x_i	$\begin{cases} \epsilon_i(p+1) \\ -\epsilon_i(p+1) \end{cases}$	$\epsilon_i \theta_i$	0

TABLE VI.

Similarly the relation (*) for $\sigma = \rho$ is written as

$$1 + (3 - u - v)(3 - \bar{u} - \bar{v}) + u\bar{u} + v\bar{v} + (1 - u)(1 - \bar{u}) + (1 - v)(1 - \bar{v}) + \sum_i (\theta_i + 2 - u - v)(\theta_i + 2 - \bar{u} - \bar{v}) = 4 + \sum_i \theta_i\bar{\theta}_i$$

where $u = \epsilon_2 X_2(\rho)$, $v = \epsilon_3 X_3(\rho)$ and the bar indicates the conjugate complex numbers. Using the relation $\sum_i \theta_i = 0$, we get

$$(2^{\alpha-1} + 1)(u + v - 2)(\bar{u} + \bar{v} - 2) + (u - v)(\bar{u} - \bar{v}) = 0,$$

or $u = v = 1$. Hence we obtain the second column of table VI.

For the last column, we get

$$1 + (1 + w + z)(1 + \bar{w} + \bar{z}) + w\bar{w} + z\bar{z} + (1 - w)(1 - \bar{w}) + (1 - z)(1 - \bar{z}) + (2^{\alpha-1} - 1)(w + z)(\bar{w} + \bar{z}) = 4,$$

setting $w = \epsilon_2 X_2(v)$ and $z = \epsilon_3 X_3(v)$. Hence

$$2^{\alpha-1}(w + z)(\bar{w} + \bar{z}) + 2w\bar{w} + 2z\bar{z} = 0,$$

or $w = z = 0$. This completes the last column of table VI.

Since N is isomorphic with $SL(2, p)$, N contains only one element of order 2. Hence there is no element of order 2 which transforms ρ or v into its inverse. This means that neither ρ nor v appears in the expansion $\langle \tau \rangle^2$. Hence by Lemma 3 we obtain two relations:

$$1 + p^2\epsilon_1(\mathcal{D}_g X_1)^{-1} + (\tfrac{1}{2}(p + 1))^2\epsilon_2(\mathcal{D}_g X_2)^{-1} + (\tfrac{1}{2}(p + 1))^2\epsilon_3(\mathcal{D}_g X_3)^{-1} = 0,$$

and

$$1 - p^2\epsilon_1(\mathcal{D}_g X_1)^{-1} + (\tfrac{1}{2}(p - 1))^2\epsilon_4(\mathcal{D}_g X_4)^{-1} + (\tfrac{1}{2}(p - 1))^2\epsilon_5(\mathcal{D}_g X_5)^{-1} = 0.$$

Note that characters belonging to a block B of defect 1 or α do not contribute to these sums, since $\sum_{s \in B} s(\sigma) = 0$ for all 2-singular elements σ. We obtain moreover the following degree relations from Table V.

$$1 + \epsilon_1\mathcal{D}_g X_1 + \epsilon_2\mathcal{D}_g X_2 + \epsilon_3\mathcal{D}_g X_3 = 0$$

$$1 + \epsilon_2\mathcal{D}_g X_2 + \epsilon_4\mathcal{D}_g X_4 = 0$$

$$1 + \epsilon_3\mathcal{D}_g X_3 + \epsilon_5\mathcal{D}_g X_5 = 0.$$

We can solve these equations for $\epsilon_i\mathcal{D}_g X_i$. Put $\xi_i = \epsilon_i\mathcal{D}_g X_i$ $(i = 1, 2, \cdots, 5)$. Then the first two relations are

$$\xi_1\xi_2\xi_3 + p^2\xi_2\xi_3 + [\tfrac{1}{2}(p + 1)]^2\xi_1(\xi_2 + \xi_3) = 0,$$

$$\xi_1\xi_4\xi_5 - p^2\xi_4\xi_5 + [\tfrac{1}{2}(p - 1)]^2\xi_1(\xi_4 + \xi_5) = 0.$$

We use the relations $-\xi_4 = 1 + \xi_2$, $-\xi_5 = 1 + \xi_3$ and make these substitutions in the second equation. Then we get

$$\xi_1(1+\xi_2)(1+\xi_3) - p^2(1+\xi_2)(1+\xi_3) - [\tfrac{1}{2}(p-1)]^2\xi_1(2+\xi_2+\xi_3) = 0.$$

Now put $X = \xi_1$ and $Y = \xi_2\xi_3$; then $\xi_2 + \xi_3 = -(1+X)$. We have then two equations for X and Y, i. e.

$$XY + p^2Y - [\tfrac{1}{2}(p+1)]^2X(1+X) = 0,$$

$$X(Y-X) - p^2(Y-X) - [\tfrac{1}{2}(p-1)]^2(1-X)X = 0.$$

Add these two equations and then divide by X. We get

$$2Y = X(p+1) - \tfrac{1}{2}(p^2-1).$$

Substituting this value in the first equation, we obtain

$$(X+p^2)\tfrac{1}{2}(p+1)^2(X - \tfrac{1}{2}(p-1)) - [\tfrac{1}{2}(p+1)]^2(1+X)X = 0,$$

or $X^2 - 2pX + p^2 = 0$, and hence $X = p$.

Since $X = \xi_1 = \epsilon_1 \mathfrak{D}_g X_1$, we must have $\epsilon_1 = 1$ and $\mathfrak{D}_g X_1 = p$. From table VI, we see that $X_1(\tau) = p$. This means that τ is contained in the kernel K of the representation corresponding to X_1. K is a normal subgroup of G, and hence $K \cap N$ is a normal subgroup of N, which contains τ but not ρ. Since $N \cong SL(2,p)$, the only normal subgroup of N which contains τ but not ρ is the center $T = \{\tau\}$. Hence $K \cap N = T$. By assumption N is the normalizer of τ in G. Hence the normalizer of τ in K coincides with T. By a theorem of Burnside, K contains a normal subgroup G_0 such that $G_0T = K$ and $G_0 \cap T = e$. If σ is any element $\neq 1$ of G_0, σ does not commute with τ, so G_0 admits an automorphism of order 2 which leaves only the identity invariant. By a lemma of Zassenhaus G_0 must be abelian.

We have to prove that the factor group G/G_0 is isomorphic with $SI(2,p)$. $K = G_0T$ is a normal subgroup of G, and T is a 2-Sylow subgroup of K, since G_0 is of odd order. Hence G is a join of K and the normalizer of T in G. But N is the normalizer of T, so that $G = KN$, or $G = G_0N$. Hence $G/G_0 \cong N/G_0 \cap N = N$.

Let P be any Sylow subgroup of G_0. If P is cyclic and $\neq e$, then the centralizer Q of P in G is a normal subgroup with cyclic factor group. $Q \cap N$ is then a normal subgroup of N, but it does not contain T. Hence $Q \cap N = e$; but this is impossible because N is not cyclic. Hence every Sylow subgroup of G_0 is not cyclic. This completes the proof of Theorem D for the case where $p \equiv 1 \pmod 4$.

If $p \equiv 3 \pmod 4$, we can use the same method by changing the signs suitably. Hence our Theorem D, and therefore Theorem C also, is proved.

It can be proved that Theorem D is still true if we assume that the normalizer of an element of order 2 is isomorphic with $SL(2, q)$ over any finite field with prime power $q = p^\mu$ elements $(p > 2)$. The proof is quite similar to the proof given here, and depends upon those properties of irreducible characters of $SL(2, q)$ which are given by Jordan [9] and Schur [10].

7. We shall add a few remarks. It is well-known that if every abelian subgroup of a finite group G is cyclic, then Sylow subgroups of G are either cyclic or a generalized quaternion group. If moreover G is non-solvable, then a 2-Sylow subgroup must be a generalized quaternion group. Thus our Theorem C gives the structure of such groups. The complete characterization may be found in Theorem E below.

There are many occasions that we have a finite group G in which every abelian subgroup is cyclic. We shall mention here two cases: (a) G is a multiplicative group contained in some non-commutative field, and (b) the cohomology groups of G over the rational integers are periodic. In both cases every abelian subgroup of G is cyclic. (For the latter case (b) cf. the recent work by Artin and Tate.)

Finally we shall consider the converse to our main theorem. Let G be a finite non-solvable group, all of whose abelian subgroups are cyclic. Then by Theorem C, G contains a normal subgroup G_1 such that $[G : G_1] \leqq 2$ and $G_1 = Z \times L$, where Z is a Z-group and $L \cong SL(2, p)$. But a group having this structure may contain an abelian subgroup which is not cyclic. For instance, the group consisting of all 2×2 matrices over the prime field with determinant $+1$ or -1 has the above structure, i.e. it contains $SL(2, p)$ as a subgroup of index 2, but it contains non-cyclic subgroups of order 4. If G_1 coincides with G, then a necessary and sufficient condition is simply that the order of Z is relatively prime to that of L. However if G_1 is a proper subgroup of G, the condition is not so simple. To state the condition explicitly we define an automorphism θ of $SL(2, p)$ by

$$\theta : \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ -\omega & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \to \begin{pmatrix} 0 & \omega^{-1} \\ -\omega & 0 \end{pmatrix}$$

where ω is a primitive element, i.e. a generator of the multiplicative group of the underlying field. In fact, θ is induced in $SL(2, p)$ by the matrix transformation by $\begin{pmatrix} 0 & -1 \\ \omega & 0 \end{pmatrix}$. It is easy to see that $\theta^2 = 1$ and θ is not an

inner automorphism of $SL(2,p)$. Moreover if f is any automorphism of $SL(2,p)$ which is not inner, then f induces an automorphism $\bar{f}$ of the central factor group $LF(2,p)$. If $\bar{f}$ were inner, then $SL(2,p)$ would admit a non-trivial automorphism of the form $\lambda \rightarrow \lambda\tau(\lambda)$ where $\tau(\lambda)$ is a central element. Since $\tau(\lambda)$ would be a homomorphism of $SL(2,p)$ into its center, $\tau(\lambda)$ would be trivial. Hence $\bar{f}$ is not inner. By Lemma 2, $\theta^{-1}f$ induces an inner automorphism of the central factor group. Using the same argument as above, we conclude that $\theta^{-1}f$ is an inner automorphism of $SL(2,p)$. This θ has the following property.

LEMMA 9. *If $\lambda\theta(\lambda)$ is in the center of L for some $\lambda \in L$, then $\lambda\theta(\lambda) = 1$, where $L \cong SL(2,p)$.*

Proof. We may use the matrix notation. Let $\lambda = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, $\alpha\delta - \beta\gamma = 1$. Then $\lambda\theta(\lambda)$ is in the center if and only if

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\begin{pmatrix} 0 & -1 \\ \omega & 0 \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\begin{pmatrix} 0 & \omega^{-1} \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon \end{pmatrix},$$

where $\epsilon = \pm 1$. A little computation shows that $\epsilon = 1$.

Using this θ, we now state the final theorem.

THEOREM E. *Every abelian subgroup of a non-solvable group G of finite order is cyclic if and only if either* (a) *G is a direct product of a Z-group Z and a group isomorphic with $SL(2,p)$ for some $p > 3$ and having order relatively prime to the order of Z, or* (b) *G is generated by $L \cong SL(2,p)$ and three elements, σ, ρ and τ such that*

$$\sigma^m = \rho^n = \tau^4 = 1,$$

$$\rho\sigma\rho^{-1} = \sigma^r, \qquad \tau\sigma\tau^{-1} = \sigma^{-1}, \qquad \tau\rho = \rho\tau$$

$$\rho\lambda = \lambda\rho, \qquad \sigma\lambda = \lambda\sigma, \qquad \tau\lambda\tau^{-1} = \theta(\lambda)$$

for any $\lambda \in L$, and τ^2 is in the center of L. Here m, n, and r are integers such that

$$(m,n) = 1, \quad (mn, p(p^2-1)) = 1 \text{ and } r^n \equiv 1 \pmod{m}.$$

Proof. Suppose that every abelian subgroup of G is cyclic. Then G contains a subgroup G_1 such that $[G:G_1] \leqq 2$ and $G_1 = Z \times L$ where Z is a Z-group and $L \cong SL(2,p)$. Since all abelian subgroups are cyclic, the order of Z must be prime to $p(p^2-1)$. If $G_1 = G$, then we are finished.

Assume $G \neq G_1$. Then G contains a subgroup L_0 such that L is a subgroup of L_0 with index 2. Since the 2-Sylow subgroups of L_0 are generalized quaternion groups, L_0 contains an element π such that $\pi \notin L$ and π^2 is in the center of L. π induces an automorphism of L, which is not inner. Hence there exists an element $\lambda_0 \in L$ such that $\pi\lambda\pi^{-1} = \lambda_0\theta(\lambda)\lambda_0^{-1}$ for every $\lambda \in L$. Since π^2 is in the center of L, we must have

$$\lambda = \pi^2\lambda\pi^{-2} = \lambda_0\theta(\lambda_0)\lambda(\lambda_0\theta(\lambda_0))^{-1},$$

or $\lambda_0\theta(\lambda_0)$ is an element of the center of L. By Lemma 9, we get $\lambda_0\theta(\lambda_0) = 1$. Let $\tau = \lambda_0^{-1}\pi$. Then

$$\tau^2 = \lambda_0^{-1}\pi\lambda_0^{-1}\pi = \lambda_0^{-1}(\pi\lambda_0^{-1}\pi^{-1})\pi^2$$
$$= \lambda_0^{-1}(\lambda_0\theta(\lambda_0)^{-1}\lambda_0^{-1})\pi^2 = (\lambda_0\theta(\lambda_0))^{-1}\pi^2 = \pi^2.$$

Since G/L is a Z-group, we can take a set of generators $\{\sigma, \rho\}$ of Z which satisfies all the requirements of Theorem E.

Conversely we assume that G has the structure in this theorem. Then it is easy to see that any l-Sylow subgroup of G is cyclic if $l > 2$. It remains to show that a 2-Sylow subgroup is a generalized quaternion group. It suffices now to show that a subgroup $\{L, \tau\}$ contains only one element of order 2. Suppose $(\lambda\tau)^2 = 1$ for some $\lambda \in L$. Then $1 = \lambda\tau\lambda\tau = \lambda(\tau\lambda\tau^{-1})\tau^2 = \lambda\theta(\lambda)\tau^2$. It follows therefore that $\lambda\theta(\lambda)$ is in the center. By Lemma 9, this implies that $\lambda\theta(\lambda) = 1$, or $\tau^2 = 1$. This is clearly impossible. Hence the 2-Sylow subgroups of G are generalized quaternion groups and our assertion is proved.

UNIVERSITY OF ILLINOIS.

REFERENCES.

[1] R. Brauer, "Investigations on group characters," *Annals of Mathematics*, vol. 42 (1941), pp. 936-958.

[2] ———, "On the arithmetic in a group ring," *Proceedings of the National Academy of Science*, vol. 30 (1944), pp. 109-114.

[3] ———, "On blocks of characters of groups of finite order. I," *Proceedings of the National Academy of Science*, vol. 32 (1946), pp. 182-186.

[4] ——— and C. Nesbitt, "On the modular characters of groups," *Annals of Mathematics*, vol. 42 (1941), pp. 556-590.

[5] ——— and K. A. Fowler (to appear).

[6] ———, M. Suzuki and H. S. Wall (to appear).

[7] W. Burnside, *The Theory of Groups*, ed. 1911.

[8] O. Grün, "Beiträge zur Gruppentheorie. I," *Journal für die reine und angewandte Mathematik*, vol. 174 (1936), pp. 1-14.

[9] H. E. Jordan, "Group-characters of various types of linear groups," *American Journal of Mathematics*, vol. 29 (1907), pp. 387-405.

[10] I. Schur, "Untersuchungen über die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen," *Journal für die reine und angewandte Mathematik*, vol. 132 (1907), pp. 85-137.

[11] M. Suzuki, "A characterization of simple groups $LF(2, p)$," *Journal of the University of Tokyo*, vol. VI-4 (1951), pp. 259-293.

[12] H. Zassenhaus, "Kennzeichung endlicher linearer Gruppen als Permutationsgruppen," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 11 (1936), pp. 17-40.

[13] ———, "Über endliche Fastkörper," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 11 (1936), pp. 187-220.

[14] ———, *Lehrbuch der Gruppentheorie*, vol. I, 1937.

ASYMPTOTIC INTEGRATIONS OF ORDINARY NON-LINEAR DIFFERENTIAL EQUATIONS.*

By PHILIP HARTMAN and AUREL WINTNER.

Introduction. It is well-known that Poincaré's classical results on systems of differential equations $dy^i/dt = f^i(y^1, \cdots, y^d)$, $i = 1, \cdots, d$, where the functions f^i are power series convergent near, and vanishing at, the origin and have a non-vanishing Jacobian there, can be extended so as to dispose of the conditions of analyticity; in fact, this matter has an extensive literature. In what follows, the problem will be taken up in its full generality, by methods previously applied by us in particular cases (for the non-conservative linear cases, cf. [6]; for binary non-linear cases, cf., e. g., [3], [4]). The generality of the methods will have two different aspects. On the one hand, nothing like a Lipschitz condition will be needed (so that no local uniqueness is assumed) and, on the other hand, the matrix of the linear approximation is allowed to have characteristic numbers and elementary divisors of arbitrary multiplicities. The continuous functions f^i will be allowed to depend on t also.

As in [6], the problem of the asymptotic integration of the system near $(y^1, \cdots, y^d) = (0, \cdots, 0)$ will be treated at two different stages of refinement; first, on a "logarithmic scale" (Part I) and, second, on the basis of the resulting rough asymptotic formulae, without this logarithmic reduction (Part II). The results of the available literature on the first stage are more complete than those on the second stage.

In order to describe the problems in a more detailed form, consider the system of differential equations

$$y' = Jy + F(t, y) \tag{1}$$

(with $y' = dy/dt$), where J is a constant d by d matrix, y and F are d-dimensional (column) vectors, and $F(t, y)$ is defined and "small" for large t and small $|y|$. The method to be employed is a combination of the procedure involving Riccati differential equations as majorants (in the same way as in the linear cases treated in [2], [6]) and of the procedure involving "zero velocity curves" (as in the binary cases of (1) treated in [3]). The variant of the "zero velocity curve" argument (cf. Sections 7, 13 below)

---

* Received January 19, 1955.

replaces that involving the superposition principle in Sections 7 and 16 of [6].

For the sake of simplicity, only the case where the elements of J, F and y are complex-valued will be treated. Hence there is no loss of generality in supposing that J is in its Jordan normal form, say, that the linear system

$$y' = Jy \tag{2}$$

has the form

$$y^{j1\prime} = \lambda y^{j1}, \tag{3_{j1}}$$

$$y^{jk\prime} = \lambda y^{jk} + y^{j\,k-1}, \qquad k = 2, \cdots, h, \tag{3_{jk}}$$

where $\lambda = \lambda(j)$; $h = h(j)$; $j = 1, \cdots, g$; $h(1) + \cdots + h(g) = d$; and (3_{jk}), for $k > 1$, is missing if $h(j) = 1$. Corresponding to the notation y^{jk}, where $j = 1, \cdots, g$ and $k = 1, \cdots, h(j)$, for the d components of y, let $F^{jk} = F^{jk}(t, y)$ denote the corresponding d components of $F(t, y)$. Thus (1) has the form

$$y^{j1\prime} = \lambda y^{j1} + F^{j1}(t, y), \tag{4_{j1}}$$

$$y^{jk\prime} = \lambda y^{jk} + y^{j\,k-1} + F^{jk}(t, y), \qquad k = 2, \cdots, h. \tag{4_{jk}}$$

It will be clear from the proofs that the methods to be used supply analogous theorems for the case in which the elements of J, F and y are real-valued; so that the normal form of J involves binary blocks of the type

$$\begin{pmatrix} \mu & \nu \\ -\nu & \mu \end{pmatrix}.$$

Part I of the paper deals with the asymptotic integration of (1) on a logarithmic scale. Theorems (*), (**) below are analogues of Theorem (*) of [6] in the linear cases of (1) and of Satz 11 of Perron [11], pp. 148-149 (cf. [12]), dealing with non-linear systems. In Theorems (*) and (**) below, the condition on $F(t, y)$ is that $|F(t, y)| / |y| \to 0$ as $(t, y) \to (\infty, 0)$ (or even relaxed forms of this condition; cf. Section 10). Perron [11], p. 149, requires a Lipschitz condition of the type

$$|F(t, y_1) - F(t, y_2)| / |y_1 - y_2| \to 0 \text{ as } (t, y_1, y_2) \to (\infty, 0, 0),$$

and his assertion is less sharp than the assertions (11), (12) of (*), (**) below. On the other hand, Perron's theorem contains an assertion of uniqueness. Such an assertion in (*) is impossible, since the conditions on $F(t, y)$ do not even imply the local uniqueness of solutions of (1). (See Section 2, for references to known cases of (*), (**), with $d = 2$.)

Part II deals with asymptotic formulae for solutions $y = y(t)$ of (1),

rather than with asymptotic formulae for the logarithms of their magnitudes. The theorems (†), (††) below are analogues of the refinement (**) in [6] of Dunkel's theorem [1] dealing with linear systems. The theorems (†), (††) are known in cases of conservative binary ($d=2$) systems. For references to the literature, see Section 12 below.

Part I. Solutions tending to the origin.

1. Notation. Let $\lambda(1),\cdots,\lambda(g)$ denote the eigenvalues of J belonging to its g elementary divisors; cf. (3). It is not assumed that the numbers $\lambda(1),\cdots,\lambda(g)$ are distinct. Let $\mu(j)=\operatorname{Re}\lambda(j)$ and let $\mu^1<\cdots<\mu^f$ denote the distinct numbers among the g numbers $\mu(1),\cdots,\mu(g)$. For a fixed m, where $1\leqq m\leqq f$, an integer j in the set $1,2,\cdots,g$ will be denoted by p, q or r according as $\mu(j)<\mu^m$, $\mu(j)=\mu^m$ or $\mu(j)>\mu^m$:

$$\mu(j)\lessgtr\mu^m \text{ according as } j=p,\ q \text{ or } r. \tag{5}$$

Put

$$L_m=\sum_q\sum_{k=1}^{h(q)}|y^{qk}|^2; \tag{6}$$

so that $|y|^2=L_1+\cdots+L_f$ is the square of the length of the vector y.

From time to time, it will be convenient to use the conventional notation $y=(y^1,\cdots,y^d)$ for the components of y and to consider the set, say $\lambda_1,\cdots,\lambda_d$, of the d eigenvalues of J. In such cases, it will be assumed that the numeration is chosen so that the k-th component y^k corresponds to the k-th eigenvalue λ_k, in the sense that the differential equation for y^k in (1) or (2) is of the form $y^{k\prime}=\lambda_k y^k+\cdots$.

2. Theorems (*) and (**). In terms of the above notations, the following theorem will first be proved:

(*) *Let the linear system of differential equations* (2) *with* (*complex-valued*) *constant coefficients have the form* (3). *Let* $F(t,y)$ *be a vector* (*with complex-valued components*) *defined and continuous for* $0\leqq t<\infty$, $|y|\leqq$ const. *and let*

$$|F(t,y)|/|y|\to 0 \text{ as } (t,y)\to(\infty,0). \tag{7}$$

Let the number $\mu=\mu^m$ *be negative,*

$$\mu^m<0. \tag{8}$$

Then, if η is any number on the range $0<\eta<1$, there exist positive numbers $\delta=\delta(\eta)$ and $T=T(\eta)$ with the property that if $t_0 \geqq T$ and if the set of $\sum h(p)+\sum h(q)$ numbers y_0^{pk} and y_0^{qk} satisfies

$$(9_1)\qquad \sum_p \sum_k |y_0^{pk}|^2 \leqq \eta \sum_q \sum_k |y_0^{qk}|^2, \qquad (9_2)\qquad 0<\sum_q \sum_k |y_0^{qk}|^2 < \delta,$$

the system (1) has a solution $y=y(t)$ which exists on the half-line $t_0 \leqq t<\infty$, satisfies the (partial set of) initial conditions

$$(10_1)\qquad y^{pk}(t_0)=y_0^{kp}, \qquad (10_2)\qquad y^{qk}(t_0)=y_0^{qk},$$

and, as $t\to\infty$, the limit relations

$$(11)\quad L_j=o(L_m) \text{ if } j\neq m, \qquad (12)\quad \log|y(t)|=(\mu^m+o(1))t.$$

It is understood that if $m=1$, so that $\mu^m=\min(\mu(1),\cdots,\mu(g))$ and the set of integers p is vacuous, then (9_1) and (10_1) are missing. If $m=f$, so that $\mu^m=\max(\mu(1),\cdots,\mu(g))$ and the set of integers r is vacuous, then (10_1)-(10_2) constitutes a complete set of initial conditions.

The cases of (*) in which it is assumed that $h(q)=1$ (that is, that the elementary divisors belonging to the $\lambda(k)$ satisfying $\operatorname{Re}\lambda(k)=\mu^m$ are simple) are contained in [15], pp. 293-294, and [13], pp. 25-26, under the additional assumption of local uniqueness.

A partial converse of (*) is as follows:

(**) *If $F(t,y)$ is a continuous vector (with complex-valued components) for $0\leqq t<\infty$, $|y|\leqq$ const. satisfying (7) and if $y=y(t)$ $(\not\equiv 0)$ is a solution of (1) for large t such that*

$$(13)\qquad y(t)\to 0 \text{ as } t\to\infty,$$

then there is an integer m, $1\leqq m\leqq f$, with the properties that $\mu=\mu^m$ satisfies $\mu\leqq 0$ and that (11) and (12) hold.

In order to illustrate the meaning of the assertion (11), suppose that $y=(y^1,y^2)$ is a binary vector, that $\mu_1\neq\mu_2$, where $\mu_k=\operatorname{Re}\lambda_k$, and that $\mu_1<0$. Then, by (*), the system (1) has solutions $y(t)=(y^1(t),y^2(t))$ satisfying (13) and

$$(14)\qquad y^2(t)/y^1(t)\to 0 \text{ as } t\to\infty.$$

According to (**), if $\mu_2>0$, then all solutions $y=y(t)$ $(\not\equiv 0)$ of (1) satisfying (13) also satisfy (14). But if $(\mu_1\neq)\ \mu_2\leqq 0$, then all solutions $y(t)$ $(\not\equiv 0)$ satisfying (13) satisfy either (14) or

$$(15)\qquad y^1(t)/y^2(t)\to 0 \text{ as } t\to\infty.$$

No positive assertion of the type (14) or (15) is possible when $\mu_1 = \mu_2$ (< 0), without further restrictions on F; cf. [5], p. 499. The conservative binary cases of (*) and (**) are contained in [7], [8], [3], [4].

3. A lemma. Lemmas 1 bis and 2 bis of Section 3 in [6], dealing with Riccati differential inequalities, will be needed below. In addition, there will be needed the following O-variant of Lemma 1 bis:

LEMMA. *Let $c > 0$, $0 < c_0 < \frac{1}{2}$ and let c^0 denote a number satisfying*

$$0 < c^0 < c_0 \text{ and } c^0 < \min cv(1-v) \text{ for } c_0 \leqq v \leqq 2c_0. \tag{16}$$

Let $f = f(t)$ be a real-valued, continuous function on $T \leqq t < \infty$ satisfying

$$f(t) \leqq c^0, \qquad (T \leqq t < \infty), \tag{17}$$

or, more generally,

$$\int_U^{U+V} f(t)dt \leqq c^0(1+V) \text{ for } T \leqq U < U+V. \tag{17 bis}$$

Finally, let $v = v(t)$ be a real-valued, continuously differentiable function on some interval $(T \leqq)$ $t_0 \leqq t \leqq t_1$, satisfying

$$v' \geqq cv(1-v) - f(t) \tag{18}$$

and

$$v(t_0) \geqq 2c_0. \tag{19}$$

Then, for $t_0 \leqq t \leqq t_1$,

$$v(t) > c_0. \tag{20}$$

In the case (17), the inequality (20) can trivially be improved to $v(t) \geqq 2c_0$ since $v'(t) > 0$ whenever $v(t) = 2c_0$.

In order to prove the assertion under the assumption (17 bis), suppose that (20) fails to hold for some t, where $t_0 \leqq t \leqq t_1$. Then $[t_0, t_1]$ contains an interval $(U, U+V)$ such that $v(U) = 2c_0$, $v(U+V) = c_0$ and $c_0 \leqq v(t) \leqq 2c_0$ *on* $(U, U+V)$. A quadrature of (18) and the second part of (16) give

$$v(U+V) - v(U) > c^0 V - \int_U^{U+V} f(t)dt.$$

The left side of this inequality is $c_0 - 2c_0 = -c_0 < -c^0$. Hence

$$\int_U^{U+V} f(t)\,dt > c^0(1+V).$$

Since this contradicts (17 bis), the Lemma follows.

4. Proof of (*). Preliminaries. Let the number a, where $0 < a < 1$, be chosen so small that

$$c = c_m = 2(\mu^m - \mu^{m-1} - 2a) > 0 \quad \text{for } m = 2, \cdots, f. \tag{21}$$

Put

$$L = L_{ma} = \sum_q \sum_{k=1}^{h(q)} a^{2k} \mid y^{qk} \mid^2, \tag{22}$$

so that

$$a^{-2}L_{ma} \leqq L_m \leqq a^{-2d}L_{ma}. \tag{23}$$

For $1 \leqq m \leqq f$, put

$$M = M_m = \sum_{j=1}^{m-1} L_{ja} \text{ and } N = N_m = \sum_{j=m}^{f} L_{ja}. \tag{24}$$

Then, for a solution $y = y(t)$ of (1) on some interval, the expressions (22), (24) become functions of t. The function (22) then satisfies

$$\mid L_{ma}' - 2\mu^m L_{ma} \mid \leqq 2aL_{ma} + d \mid F(t,y) \mid \cdot \mid y \mid, \tag{25_m}$$

and so

$$M' \leqq 2(\mu^{m-1} + a)M + d^2 \mid F(t,y) \mid \cdot \mid y \mid, \tag{26_m}$$

$$N' \geqq 2(\mu^m - a)N - d^2 \mid F(t,y) \mid \cdot \mid y \mid; \tag{27_m}$$

cf. Section 4 in [6].

Although the conditions of (*) on (1) do not imply the unique determination of solutions of (1) by their initial conditions, any solution $y = y(t)$ vanishing at some large value of t, say $t \geqq T_0$, is identically 0. This is clear from condition (7) of (*). Hence, if $y = y(t) \not\equiv 0$ is a solution of (1) on some interval $U \leqq t \leqq U + V$, where $U \geqq T_0$, then $y(t) \neq 0$. Hence $M + N \geqq a^{2d} \mid y \mid^2 > 0$. Consequently, it is possible to define an auxiliary function $v = v_m(t)$ by

$$v = N_m/(N_m + M_m), \tag{28}$$

and this function satisfies

$$0 \leqq v \leqq 1 \tag{29}$$

and

$$v' \geqq c_m v(1-v) - d^2 a^{-2d} |F(t,y)|/|y|; \tag{30}$$

cf. Section 4 in [6].

If $m < f$, then it is possible to define the function $w = w_m(t)$ $(= v_{m+1}(t))$ by

$$w = N_{m+1}/(N_m + M_m), \tag{31}$$

and this function satisfies

$$0 \leqq w \leqq 1 \tag{32}$$

and has a continuous derivative satisfying

$$w' \geqq c_{m+1} w(1-w) - d^2 a^{-2d} |F(t,y)|/|y|. \tag{33}$$

5. The case $m = f$. As in the linear case treated in [6], the case $m = f$ is quite simple. If $m = f$, the real parts μ^j of all characteristic numbers of J satisfy $\mu^j \leqq \mu < 0$. This will be shown to imply that if $a > 0$ is fixed so that $\mu + 2a < 0$, then there exist positive numbers $T = T(a)$ and $\epsilon = \epsilon(a)$ with the property that if

$$\sum_p \sum_k |y_0^{pk}|^2 + \sum_q \sum_k |y_0^{qk}|^2 < \tfrac{1}{2}\epsilon a^{2d} \tag{34}$$

and if $t_0 \geqq T$, then a solution satisfying the (complete set of) initial conditions (10_1)-(10_2) has extensions over the half-line $t_0 \leqq t < \infty$ and all such extensions satisfy

$$|y| < \epsilon \text{ for } t \geqq t_0. \tag{35}$$

In order to see this, let $T = T(a)$ and $\epsilon(a)$ be chosen so that (35) and

$$t \geqq T \tag{36}$$

imply

$$d^2 a^{-2d} |F(t,y)| \leqq 2a|y|. \tag{37}$$

The existence of such T and ϵ is clear from (7).

Since $m = f$, it follows from (24) that $|y|^2 \leqq a^{-2d} M_{m+1}$. Thus if $y = y(t)$ satisfies (35) on some t-interval $t_0 \leqq t \leqq t^0$, where $t_0 \geqq T$, then (26_{m+1}) and (37) show that

$$M_{m+1}' \leqq 2(\mu + 2a) M_{m+1}. \tag{38}$$

Hence, $M_{m+1}(t) \leqq M_{m+1}(t_0) e^{2(\mu+2a)(t-t_0)} \leqq M_{m+1}(t_0)$ on this interval, and so

$$|y(t)|^2 \leqq a^{-2d} |y(t_0)|^2 e^{2(\mu+2a)(t-t_0)}. \tag{39}$$

Let (34) hold. Consider a solution $y = y(t)$ of (1) determined by the initial conditions (10_1)-(10_2). Then, by continuity, (35) holds for t near $t_0 (\geqq T)$. By (34) and (39),

$$|y(t)|^2 \leqq \tfrac{1}{2}\epsilon e^{(\mu+2a)(t-t_0)} \leqq \tfrac{1}{2}\epsilon \tag{40}$$

for such $t (\geqq t_0)$. Consequently, any extension of $y(t)$, with increasing t, continues to satisfy (35) (in fact, (40)). Hence $y = y(t)$ has extensions over $t_0 \leqq t < \infty$ and all extensions satisfy (40). This proves the assertion concerning (34) and (35).

Let δ in (9_2) be so small that $2\delta < \frac{1}{2}\epsilon a^{2d}$, where $\epsilon = \epsilon(a)$ is the number defined in connection with (35), (36). Since $\eta < 1$, the inequality (34) follows from (9_1) and (9_2). Hence any solution $y = y(t)$ of (1) satisfying (10_1)-(10_2) has extensions over $t_0 \leqq t < \infty$ and any extension satisfies (40). Consequently,

$$|F(t, y(t))|/|y(t)| \to 0 \text{ as } t \to \infty. \tag{41}$$

Since $\eta < 1$, condition (9_1) implies that

$$v(t_0) \geqq 2c_0, \text{ where } c_0 = 1/2(1+\eta) < \tfrac{1}{2}. \tag{42}$$

Let $T = T(a, \eta)$ be so large that (35) and (36) imply

$$d^2 a^{-2d} |F(t, y)| \leqq c^0 |y|, \tag{43}$$

if c^0 is a constant satisfying (16), where c is c_m and c_0 is given in (42). Hence, by (30), (42) and the Lemma of Section 3, (20) holds for $t_0 \leqq t < \infty$. By (29), (41) and Lemma 1 bis in Section 3 of [6],

$$v(t) \to 1 \text{ as } t \to \infty. \tag{44}$$

The definition (28) of $v(t)$ and the limit relation (44) imply that, as $t \to \infty$,

$$L_{ja}(t) = o(L_{ma}t) \text{ if } j \neq m. \tag{45}$$

By virtue of (23), this proves assertion (11) of (*).

Since, by (45), $|y|^2 \leqq a^{-2d}(1 + o(1))L_{ma}$, the limit relation (41) shows that the differential inequality (25_m) can be written as

$$|(\log L_{ma})' - 2\mu^m| \leqq 2a + o(1).$$

Hence,

$$\limsup_{t\to\infty} |t^{-1} \log L_{ma} - 2\mu^m| \leqq 2a, \text{ and so } \limsup_{t\to\infty} |t^{-1} \log L_m - 2\mu^m| \leqq 2a,$$

by (23). Since $a > 0$ can be chosen arbitrarily small, (12) follows. This completes the proof of (*) if $m = f$.

6. The case $m < f$. If a solution $y = y(t)$ of (1) on some t-interval has an auxiliary function $v = v(t)$ satisfying (20) and an auxiliary function $w = w(t)$ satisfying

$$w(t) < \tfrac{1}{2}c_0 \tag{46}$$

for some constant $c_0 > 0$, then the definitions (28), (31) of v, w show that

$$L_{ma} > \tfrac{1}{2}c_0(M_m + L_{ma} + N_{m+1}); \tag{47}$$

and since $|y|^2 \leqq a^{-2d}(M_m + L_{ma} + N_{m+1})$,

$$|y|^2 \leqq 2a^{-2d}c_0^{-1}L_{ma}. \tag{48}$$

Thus, when (20) and (46) hold, the differential inequality (25_m) shows that

$$|L_{ma}' - 2\mu L_{ma}| \leqq \{2a + 2da^{-2d}c_0^{-1}|F(t,y)|/|y|\}L_{ma}. \tag{49}$$

Let $a > 0$ be so small that $\mu + 2a < 0$ and let $T = T(a,\eta)$ and $\epsilon = \epsilon(a,\eta) > 0$ be chosen so that (35) and (36) imply the inequalities

$$2da^{-2d}c_0^{-1}|F(t,y)| \leqq 2a|y| \tag{50}$$

and (43), where, in the latter, c^0 is a constant satisfying (16) in which c is c_m or c_{m+1} and c_0 is $1/2(1+\eta)$.

Thus, if (20), (46) and

$$2a^{-2d}c_0^{-1}L_{ma} < \epsilon \tag{51}$$

hold, then

$$L_{ma}' \leqq 2(\mu + 2a)L_{ma}, \qquad (\mu + 2a < 0). \tag{52}$$

On any interval, $t_0 \leqq t \leqq t^0$, on which (52) holds,

$$L_{ma}(t) \leqq L_{ma}(t_0)e^{2(\mu+2a)(t-t_0)} \leqq L_{ma}(t_0). \tag{53}$$

Consider a solution $y = y(t) \not\equiv 0$ on some t-interval, bounded by $t = t_0$ $(\geqq T)$ on the left, satisfying (42), (46) and (51) at $t = t_0$. For an extension of $y(t)$, with increasing t, there are two possibilities: either $y(t)$ exists on $t_0 \leqq t < \infty$ and satisfies (20), (46) and (51) throughout or there exists a least value of $t = t^0$ $(> t_0)$ such that $y(t)$ satisfies (20) and (51) on $t_0 \leqq t \leqq t^0$, and (46) on $t_0 \leqq t < t^0$, finally

$$w(t) = \tfrac{1}{2}c_0 \tag{54}$$

at $t = t^0$.

In order to see this, note that if (20), (46) and (51) hold on $t_0 \leqq t < t^0$, so that (35) holds by (48), then $y(t)$ has extensions over an interval containing $t_0 \leqq t < t^0$ in its interior. Thus, the assertion above is merely to

the effect that if (20), (46) and (51) hold on $t_0 \leqq t < t^0$, then (20), (51) also hold at $t = t^0$. In fact, (20), (46) and (51) imply (53) for $t_0 \leqq t < t^0$, so that (51) holds at t^0 since it holds at t_0. In view of (30), (36) and (43), the Lemma of Section 3 implies that (20) also holds on $t_0 \leqq t \leqq t^0$.

Note that the differential inequality (33) and the inequalities (20) (hence (48)) and (51) imply

$$w'(t) > 0 \text{ when } w(t) = \tfrac{1}{2}c_0 \text{ and } t \geqq T; \tag{55}$$

cf. (33) and (43).

7. Smooth $F(t, y)$. The proof of (*) will now be completed for the case that the solutions of (1) are uniquely determined (locally) by initial conditions (say, that $F(t, y)$ has continuous first derivatives with respect to $y^1, \cdots, y^d$). The proof will depend on a generalization of the "zero velocity curve" argument used for binary cases in [3]; this generalization can be considered as a particular case of the topological argument formulated in [14].

It will be shown that *to the partial set of conditions* (10_1)-(10_2), *where* $\delta > 0$ *is sufficiently small in* (9_2), *there belongs at least one set of* $\Sigma h(r)$ *numbers* y_0^{rk} *such that the solution determined by* (10_1), (10_2) *and*

$$y^{rk}(t_0) = y_0^{rk} \tag{56}$$

exists for $t_0 \leqq t < \infty$ *and satisfies* (46), *that is,*

$$N_{m+1} < \tfrac{1}{2}c_0(M_m + L_{ma} + N_{m+1}), \tag{57}$$

In other words, it will be shown that, among the set of $\Sigma h(r)$ numbers y_0^{rk} satisfying

$$(1 - \tfrac{1}{2}c_0) \sum_r \sum_k |a^k y_0^{rk}|^2 \leqq \tfrac{1}{2}c_0 \Big(\sum_p \sum_k |a^k y_0^{pk}|^2 + \sum_q \sum_k |a^k y_0^{qk}|^2\Big), \tag{58}$$

there is at least one set such that the solution of (1) determined by (10_1)-(10_2) and (56) exists for $t_0 \leqq t < \infty$ and satisfies (57) on this half-line.

For any choice of (56), condition (9_1) implies (36), hence (20). If (58) holds, then $w(t) \leqq \frac{1}{2}c_0$, and so (48) holds at $t = t_0$. Let δ in (9_2) be so small that (9_2) implies (51) at $t = t_0$.

Consider the (unique) solution determined by (10_1)-(10_2) and (56), when (58) holds. Corresponding to each choice of (56), define a number t^0 $(\leqq \infty)$ as follows: If equality holds in (58), let $t^0 = t_0$; if inequality holds in (58), let t^0 $(> t_0)$ be the least t-value $(\leqq \infty)$ at which (54) holds. According to Section 6, $y(t)$ has an extension on such an interval $t_0 \leqq t < t^0$

($\leqq \infty$) and (20), (46) and (51) hold on $t_0 \leqq t < t^0$. Furthermore, if $t^0 < \infty$, then (20), (51) and (54) are satisfied and (55) is applicable at $t = t^0$.

If $t^0 = \infty$ for some choice of (56) satisfying (58), then the italicized statement concerning (56) is proved. Suppose, if possible, that ($t_0 \leqq$) $t^0 < \infty$ for every choice of (56). Clearly, $y(t^0) \neq 0$. The mapping $(t_0, y(t_0)) \to (t^0, y(t^0))$ is one-to-one, by the local uniqueness of solutions of (1), and is continuous, by the continuous dependence of solutions of (1) on initial conditions, and by the fact that (55) is applicable at $t = t^0$ (so that $y = y(t)$ intersects, and is not tangent to, the "conical" surface $N_{m+1} - \frac{1}{2}c_0(N_m + L_{ma} + N_{m+1}) = 0$ at $t = t^0$).

The domain of definition, say D, of the mapping $(t_0, y(t_0)) \to (t^0, y(t^0))$ is the set of points with the $d+1$ coordinates $(t_0; y_0{}^{pk}, y_0{}^{qk}, y_0{}^{rk})$, where t_0, $y_0{}^{pk}$, $y_0{}^{qk}$ are fixed and (58) holds. Since this is a closed set, and since $t^0 < \infty$ and $y(t^0) \neq 0$, there exist positive constants such that $t_0 \leqq t^0 \leqq$ const. and $|y(t^0)| \geqq$ Const. > 0. Hence, the range, say R, of the mapping is contained in the product set of the t-interval $t_0 \leqq t \leqq$ const. and the "lateral surface of the frustrum of a cone":

$$(59) \quad \{t_0 \leqq t \leqq \text{const.}\} \times \{(1 - \tfrac{1}{2}c_0)N_{m+1} = \tfrac{1}{2}c_0(M_m + L_{ma}),\ 0 < c_1 \leqq M_m + L_{ma} \leqq c_2\},$$

where c_1 and c_2 are constants. Topologically, (59) can be considered as a product set $I_1 \times I_2 \times C(d_1) \times C(d_2)$, where I_1 and I_2 are closed (real, linear) intervals, $C(j)$ is the boundary of a j-dimensional sphere (that is, of a $2j$-dimensional real sphere), and $d_1 = \Sigma h(p) + \Sigma h(q)$, $d_2 = \Sigma h(r)$. Such a representation results, for example, by writing a point of the set (59) as $(t, \rho, e(d_1), e(d_2))$, where $t_0 \leqq t \leqq$ const., $c_1{}^{\frac{1}{2}} \leqq \rho \leqq c_2{}^{\frac{1}{2}}$, $(a^k y^{pk}, a^k y^{qk}) = \rho e(d_1)$, $(1 - \frac{1}{2}c_0)(a^k y^{rk}) = \frac{1}{2}c_0 \rho e(d_2)$, and $e(j)$ is an arbitrary unit vector with j (complex-valued) components.

Since the definition of t^0 includes the identity $t^0 = t_0$ when equality holds in (58), the set $C = C(d_2)$ of points $(t; y^{pk}, y^{qk}, y^{rk})$, where $t = t_0$, $y^{pk} = y_0{}^{pk}$, $y^{qk} = y_0{}^{qk}$ and equality holds in (58), is both in the domain D and in the range R of the mapping $(t_0, y(t_0)) \to (y^0, y(t^0))$. Consider a deformation of C and D into the point $(t_0, y_0{}^{pk}, y_0{}^{qk}, 0)$. The topological mapping $(t_0, y(t_0)) \to (t^0, y(t^0))$ induces a deformation of C in R (hence in (59)) into a point. In view of the product set structure of (59), a projection leads to a deformation of C in itself into a point. Since this is impossible, the assumption that $t^0 < \infty$ holds for all points of D is untenable. This proves the italicized statement concerning (56).

The proof of (*) for the case at hand can now be completed at once.

Let $y = y(t)$ be a solution determined by (10_1), (10_2) and (56), where (56) is chosen so that $y(t)$ exists on $t_0 \leqq t < \infty$ and satisfies (46) there. Then (20) (and (51)) holds for $t_0 \leqq t < \infty$. In particular, (35), and so (41), holds.

In view of (20), (30) and (41), Lemma 1 bis in Section 3 of [6] implies (44). Similarly, it follows from (33), (46) and (41) that

$$w(t) \to 0 \text{ as } t \to \infty. \tag{60}$$

The definitions (22), (28), (31) of L_{ma}, v, w show that (44) and (60) are equivalent to (45). The passage from (45) to (11) and (12) is the same as in the case $m = f$ treated in Section 5 above.

This proves (*) provided that the solutions of (1) are uniquely determined (locally) by initial conditions.

8. Completion of the proof of (*). In order to prove (*) without this proviso, note that if t^0 $(> T)$ is fixed, then there exists a sequence of (vector) functions $F_1(t, y)$, $F_2(t, y), \cdots$ which are defined and smooth (say, of class C^1) for $T \leqq t \leqq t^0$, $|y| \leqq$ const. and such that

$$F_n(t, y) \to F(t, y) \text{ as } n \to \infty, \tag{61}$$

uniformly for $T \leqq t \leqq t^0$, $|y| \leqq$ const. It can also be supposed that if y and t $(\leqq t^0)$ satisfy (35) and (36) respectively, then (43_n) and (50_n) hold, where (43_n) and (50_n) result from (43) and (50), respectively, by writing F_n in place of F. It can then be supposed that the domain of definition of $F_n(t, y)$ has been extended over $T \leqq t < \infty$, $|y| \leqq$ const., so that $F_n(t, y)$ is smooth and that (43_n) and (50_n) hold whenever (35) holds (whether or not $t \leqq t^0$).

After letting $t^0 = T + 1, T + 2, \cdots$ and applying a diagonal process, it can be supposed that the sequence $F_1(t, y). F_2(t, y), \cdots$ is independent of t^0. Thus (61) holds uniformly on every set $T \leqq t \leqq$ Const., $|y| \leqq$ const., for arbitrary Const. $(> T)$. It can also be supposed that (35) and (36) imply (43_n) and (50_n) for *every n*.

According to the considerations of Section 7, the system of differential equations

$$y' = Jy + F_n(t, y) \tag{62}$$

has a solution $y = y_n(t)$ on $(T \leqq)$ $t_0 \leqq t < \infty$ satisfying (10_1)-(10_2), (20), (46) and (51). By virtue of (46), the functions M_m, L_{ma}, N_{m+1} belonging to $y_n(t)$ satisfy $(1 - \frac{1}{2}c_0)N_{m+1} < \frac{1}{2}c_0(L_{ma} + M_m)$. Since $\frac{1}{2}c_0 < 1$ and since L_{ma},

M_m are independent of n at $t=t_0$, by (10_1)-(10_2), it follows that the components $y^{rk}(t_0)=y_n{}^{rk}(t_0)$ of $y_n(t_0)$ are bounded (as functions of n). Hence there exists a subsequence of the integers $1, 2, \cdots$ (which can be supposed to be the full sequence after a renumeration) such that

$$\lim_{n\to\infty} y_n(t)=y(t) \text{ exists} \tag{63}$$

at $t=t_0$. (In view of (10_1)-(10_2), this merely requires the existence of the limits

$$\lim_{n\to\infty} y_n{}^{rk}(t_0)=y_0{}^{rk}$$

for every (rk).)

The uniformity of the limit process (61) implies that, after another selection, it can be supposed that (63) holds uniformly on every bounded t-interval $t_0\leqq t\leqq$ Const. and that the limit $y=y(t)$ is a solution of (1) on $t_0\leqq t<\infty$.

The solution $y=y(t)$ of (1) satisfies (10_1)-(10_2) since $y=y_n(t)$ does. In view of (63), $y(t)$ satisfies the inequalities which result from (20), (46) and (51) if $<$ is replaced by $\leqq$. Finally, it can be verified as in Section 5 that $y=y(t)$ satisfies (11) and (12). This completes the proof of (*).

9. Proof of (**). The function $v(t)=v_m(t)$ belonging to the solution $y=y(t)$ in the statement of (**) satisfies (29) and (30) for large t, where (7) and (13) imply (41). Lemma 1 bis in Section 3 of [6] shows that

$$(64_1)\quad \lim_{t\to\infty} v_m(t)=1 \qquad \text{or} \qquad (64_2)\quad \lim_{t\to\infty} v_m(t)=0$$

for $m=1, 2, \cdots, f$. It is clear from the definitions (24) and (28), that if (64_2) holds for some $m=m_0$, then (64_2) holds for $m\geqq m_0$. Also, (64_1) holds if $m=1$, since $v_1(t)\equiv 1$. Let m denote the largest integer (with $1\leqq m\leqq f$) for which (64_1) holds.

Thus $v(t)=v_m(t)$ satisfies (44) and, if $m<f$, $w(t)=v_{m+1}(t)$ satisfies (60). These relations imply (45) and, therefore, (11) and (12); cf. the end of Section 5. Finally, $\mu\leqq 0$ follows from (12) and (13). This proves (**).

10. Variants of (*). It will be shown that the condition (7) in (*) can be replaced by an "averaged" form of (7), a condition of the type employed in [2] and in the linear case [6] of (1). The methods of proof of this fact will have applications below.

(*bis) *Assertion* (*) *remains valid if* (9_2) *is replaced by*

$$e^{2(\mu-\delta)t}<\sum_q\sum_k |\, y_0{}^{qk}\,|^2<e^{2(\mu+\delta)t} \tag{65}$$

and if condition (7) *on* $F(t,y)$ *is replaced by the requirement that, for sufficiently small* $\delta > 0$, *there exists a continuous function* $f(t) = f(t;\delta)$ *with the properties that*

$$|F(t,y)| \leqq |y| f(t) \text{ if } e^{(\mu-\delta)t} < |y| < e^{(\mu+\delta)t} \tag{66}$$

and

$$\text{l. u. b.}_{0 \leqq V < \infty} \int_U^{U+V} f(t)\,dt/(1+V) \to 0 \text{ as } U \to \infty. \tag{67}$$

The changes in the proof of (*) needed to prove (*bis) will be indicated. It will be seen that difficulties arise from the fact that (55) is no longer valid, but that these difficulties can be avoided.

As above, let a denote a small positive number. In view of (66) and (67), there exists a $T = T(a,\eta)$ such that

$$e^{(\mu-2a)t} < |y| < e^{(\mu+2a)t} \tag{68}$$

implies the inequality in (66),

$$2d^2a^{-2d}c_0^{-1} \int_U^{U+V} f(t)\,dt \leqq 2a + 2aV \text{ if } T \leqq U \leqq U+V \tag{69}$$

and

$$d^2a^{-2d} \int_U^{U+V} f(t)\,dt \leqq c^0(1+V) \text{ if } T \leqq U \leqq U+V, \tag{70}$$

where c^0 satisfies the inequalities in (16) for $c = c_m$ and $c = c_{m+1}$ and for $c_0 = 1/2(1+\eta)$, as in (42), or if c_0 is replaced by $c_0/4 = 1/8(1+\eta)$.

Note that the applications of the Lemma in the proof of (*) remain valid if (43) is replaced by (70) (that is, if (17) is replaced by (17 bis)).

For brevity, consider the proof of (*bis) only in the case $m < f$. If (20) and (46) hold, then (47) is valid and (48) can be changed to

$$L_{ma} \leqq |y|^2 \leqq 2a^{-2d}c_0^{-1}L_{ma}. \tag{71}$$

The inequality (49) can be written in the form

$$|L_{ma}' - 2\mu L_{ma}| \leqq \{2a + 2da^{-2d}c_0^{-1}f(t)\}L_{ma} \tag{72}$$

if $t \geqq T$ and (67) hold. In view of (70), the relations (68) hold if

$$e^{2(\mu-2a)t} < L_{ma} < \tfrac{1}{2}a^{2d}c_0^{-1}e^{2(\mu-2a)t}. \tag{73}$$

When (73), hence (72), is valid on an interval $t_0 \leqq t < t^0$, a quadrature and (69) give

$$e^{-2a}e^{2(\mu-2a)(t-t_0)} \leqq L_{ma}(t)/L_{ma}(t_0) \leqq e^{2a}e^{2(\mu+2a)(t-t_0)}. \tag{74}$$

Hence (73) is assured at $t = t^0$ if

$$e^{2a}e^{2(\mu-2a)t} < L_{ma}(t) < e^{-2a}\left(\tfrac{1}{2}a^{2d}c_0^{-1}e^{2(\mu-2a)t}\right) \tag{75}$$

holds at $t = t_0$.

Thus, if $y = y(t)$ satisfies (42), (46) and (75) at $t = t_0$ ($\geqq T$), then (20) and (73) hold on $t_0 \leqq t \leqq t^0$ and (46) holds on $t_0 \leqq t < t^0$.

The condition (75) at $t = t_0$ is implied by (10_1)-(10_2) and the inequalities (9_1) and (65) if $\delta > 0$ is sufficiently small.

It is clear that *if* (55) *holds*, then the remainder of the proof of (*) becomes applicable to (*bis) (if the inequalities (9_2) and (51) in that proof are replaced by (65) and (73) respectively). It will be seen that the possible failure of (55) is not a serious difficulty.

Let $F_0(t, y)$ be a continuous function on the set of (t, y) satisfying inequalities $T \leqq t < \infty$, (68) and

$$N_{m+1} \leqq \tfrac{1}{2}c_0(M_m + L_{ma} + N_{m+1}). \tag{76}$$

Let F_0 have the properties that $F_0(t, y) \equiv F(t, y)$ when

$$N_{m+1} \leqq (c_0/4)(M_m + L_{ma} + N_{m+1}); \tag{77}$$

that, for other points (t, y) of the domain of definition of F_0,

$$|F_0(t, y)| \leqq |F(t, y)| + \theta, \tag{78}$$

where $\theta = \theta(a)$ is a positive number to be described below; finally, that

$$d^2a^{-2d}|F_0(t, y)| \leqq c^0|y| \text{ if } N_{m+1} = \tfrac{1}{2}c_0(M_m + L_{ma} + N_{m+1}). \tag{79}$$

In the considerations involving (68)-(75) and the definition of F_0, let a be replaced by a smaller positive number, say $\frac{1}{2}a$, and, correspondingly, $T = T(a, \eta)$ by $T(\frac{1}{2}a, \eta)$. Then, if $\theta = \theta(a) > 0$ in (78) is sufficiently small, it is clear that the remarks concerning (68)-(75) remain valid if $y = y(t)$ is a solution of

$$y' = Jy + F_0(t, y). \tag{80}$$

In addition, the inequality in (79), which is the analogue of (43), assures that (55) is applicable to the solutions $y(t)$ of (80).

Thus, the arguments of Section 7 show that, if δ is sufficiently small in (65), then there is on $t_0 \leqq t < \infty$ a solution $y = y(t)$ of (80) which satisfies

the (partial) set of initial conditions (10_1)-(10_2), where $t_0 \geqq T$, and the asymptotic relations (11)-(12) as $t \to \infty$.

In particular, the auxiliary function $w(t)$ of $y = y(t)$ satisfies (60). In view of the fact that $T = T(a, \eta)$ was chosen so large that (68) implies (70) if c^0 satisfies the inequalities in (16) when $c = c_{m+1}$ and c_0 is replaced by $c_0/4 = 1/8(1+\eta)$, it follows from (60) that $w(t) < c_0/4$ for $t_0 \leqq t < \infty$. For, according to the Lemma (and the Remark following its proof) in Section 3, if $w(t) \geqq c_0/4$ for some $t = t^0$ $(\geqq t_0 \geqq T)$, then $w(t) \geqq c_0/8$ for $t \geqq t^0$. This contradicts (60).

Hence, for the solution $y = y(t)$, the functions $M_m(t)$, $L_{ma}(t)$, $N_{m+1}(t)$ satisfy (77). But the defining properties of $F_0(t, y)$ show that this implies $F(t, y(t)) \equiv F_0(t, y(t))$, so that $y = y(t)$ is a solution of (1), as well as of (80). This proves (* bis).

Remark. For applications below, note that the proofs of (*) and (* bis) show that analogues of these assertions remain valid if the "linear part" (2) of (1) is replaced by a suitable system with coefficients depending on t or on (t, y); for example, by a system of the type

$$y^{j1\prime} = \lambda^{j1}(t, y) y^{j1}, \qquad y^{jk\prime} = \lambda^{jk}(t, y) y^{jk} + \delta^{jk}(t, y) y^{j\,k-1},$$

where $\lambda^{jk}(t, y)$, $\delta^{jk}(t, y)$ are complex-valued, continuous functions for large t and small $|y|$. Corresponding to the sets of integers p, q, r in (*), suppose that it is possible to divide the integers $j = 1, \cdots, g$ into three sets, the elements of which will be denoted by p, q, r, with the property that for some pair of constants c and a, where $c > 0$ and $0 < a < 1$,

$$\begin{aligned}
&\operatorname{Re} \lambda^{jk} < -c - a\,|\,\delta^{jk}\,| && \text{if } j = q,\\
&\operatorname{Re} \lambda^{pk} + a\,|\,\delta^{pk}\,| \leqq \operatorname{Re} \lambda^{qk} - a\,|\,\delta^{qk}\,| - c && \text{for all } p, q,\\
&\operatorname{Re} \lambda^{rk} - a\,|\,\delta^{pk}\,| \geqq \operatorname{Re} \lambda^{qk} + a\,|\,\delta^{qk}\,| + c && \text{for all } q, r,
\end{aligned}$$

where $\delta^{j1} \equiv 0$. Let L_m be defined by (6). Then (*) and (* bis) remain valid if (11) and (12) are replaced, respectively, by

$$|\,y^{jk}(t)\,|^2 = o(L_m(t)) \text{ if } j \neq q \quad \text{and} \quad \limsup_{t \to \infty} t^{-1} \log |\,y(t)\,| \leqq -c.$$

Part II. Asymptotic formulae.

11. The linear case. Let h_* denote the maximum multiplicity of the elementary divisors of J belonging to characteristic numbers $\lambda(j)$ satisfying $\operatorname{Re} \lambda(j) = \mu$; that is,

$$h_* = \max_q h(q). \tag{81}$$

Let h_0 denote a number (not necessarily an integer) satisfying

$$h_0 \geqq h_*. \tag{82}$$

Finally, let $q = q_0$ denote any fixed value of q, let k_0 be an integer on the range $1 \leqq k_0 \leqq h(q_0)$, and let $b = b(q_0, k_0)$ be defined by

$$b = h_0 - h(q_0) + k_0 \qquad (\text{so that } k_0 \leqq b \leqq h_0). \tag{83}$$

As a refinement of a theorem of Dunkel [1], it was shown in [6] that if (1) is a linear system, say $F^j(t, y) = \sum_{k=1}^{d} g_{jk}(t) y^k$, subject to

$$\int^{\infty} t^{h_0 - 1} |\, g_{jk}(t)\,| \, dt < \infty, \tag{84}$$

then (1) has a solution $y(t) = y_{q_0 k_0}(t)$ the components of which have the following asymptotic properties, as $t \to \infty$: when $q = q_0$,

$$y^{qk}(t) = o(e^{\mu t} t^{k-b}) \text{ if } 1 \leqq k < k, \tag{85_1}$$

$$y^{qk}(t) = e^{\lambda(q) t} t^{k-k_0}/(k - k_0)\,! + o(e^{\mu t} t^{k-b}) \text{ if } k_0 \leqq k \leqq h(q); \tag{85_2}$$

when $q \neq q_0$,

$$y^{qk}(t) = o(e^{\mu t} t^{k-b}) \text{ if } 1 \leqq k \leqq h(q); \tag{85_3}$$

finally, when $j = p$ or $j = r$,

$$y^{jk}(t) = o(e^{\mu t} t^{1-b}) \text{ if } 1 \leqq k \leqq h(j). \tag{85_4}$$

Furthermore, corresponding to every solution $y = y(t)$ of the linear cases of (1) satisfying (12), there exists a set of $\Sigma h(q)$ constants c^{qk}, not all 0, with the property that $y(t)$ has the following behavior as $t \to \infty$:

$$y(t) = \sum_{q_0} \sum_{k_0} c^{q_0 k_0} y_{q_0 k_0}(t), \tag{86}$$

where $y = y_{q_0 k_0}(t)$ is a vector satisfying the asymptotic formulae (85_1)-(85_4).

Note that if $h_0 \geqq 2h_* - 1$ (for example, if $h_0 \geqq 2d - 1$) in (84), then the remainder terms $o(e^{\mu t} t^{k-b})$ in (85_1)-(85_4) are $o(e^{\mu t})$.

It is clear from (86) that certain terms on the right of the asymptotic formula (86) may be redundant. Thus, for a fixed (qk), the error terms $o(e^{\mu t} t^{k-b})$ are $o(e^{\mu t} t^{k-\beta})$, where $\beta = h_0 - j_0$ and j_0 is the maximum of the differences $h(q_0) - k_0$ for which $c^{q_0 k_0} \neq 0$. Consequently, a "principal" term

$e^{\lambda(q)t}t^{k-k_0}$ of (85) contributes to (86) only if $k_0 \leqq \beta$. The sum (86) can be abbreviated to

$$y(t) = \sum_q \sum_{k=l(q)}^{k(q)} c^{qk} y_{qk}(t), \tag{87}$$

where $k = l(q)$, $k(q)$ are the least and greatest values of k (if any) satisfying

$$1 \leqq l(q) \leqq k(q) \leqq \min(h(q), h_0 - j_0) \text{ and } h(q) - l(q) \geqq j_0. \tag{88}$$

For a given q, the number of terms in the last sum is

$$k(q) - l(q) + 1 \leqq \min(j_0, h_0 - h(q)) + 1.$$

12. Statement of the theorems. In what follows, corresponding assertions will be proved for non-linear systems:

(†) *Let the linear system of differential equations* (2) *with (complex-valued) constant coefficients have the form* (3). *Let* $\mu = \mu^m$ *satisfy* (8) *and let* (81), (82) *hold. Let* j_0 *be a number satisfying* $0 \leqq j_0 < h_*$ *and* $l(q)$, $k(q)$ *the least, greatest integers, if any, satisfying* (88). *Let* c^{qk}, *where* $l(q) \leqq k \leqq k(q)$, *be given numbers, not all* 0. *Let* $F(t, y)$ *be a vector (with complex-valued components) defined and continuous for* $0 \leqq t < \infty$, $|y| \leqq$ const. *Finally, let* $\phi(t, \rho)$ *be a continuous (scalar) function defined for* $0 \leqq t < \infty$, $0 \leqq \rho \leqq$ const. *such that* F *and* ϕ *satisfy*

$$|F(t, y)| \leqq \phi(t, |y|), \tag{89}$$

$$\phi(t, \rho) \text{ is non-decreasing in } \rho \text{ (for fixed } t), \tag{90}$$

$$\int^{\infty} e^{-\mu t} t^{h_0 - j_0 - 1} \phi(t, \text{const. } e^{\mu t} t^{j_0}) dt < \infty \text{ for every const.} > 0. \tag{91}$$

Then (1) *has a solution* $y = y(t)$ *for large* t *satisfying the asymptotic formula* (87).

Conditions (89), (90), (91) are surely satisfied if there exists, for large t, a continuous function $f(t)$ satisfying

$$|F(t, y)| \leqq f(t)|y| \quad \text{and} \quad \int^{\infty} t^{h_0 - 1} f(t) dt < \infty. \tag{94}$$

In fact, $\phi(t, \rho)$ can then be chosen to be $f(t)\rho$.

In the conservative case, where $F(t, y) \equiv F(y)$ does not depend on t, conditions (89), (90), (91) are satisfied if there exists a continuous function $\phi(\rho)$ satisfying the following conditions:

$$|F(y)| \leqq \phi(|y|), \tag{95}$$

$$\phi(\rho) \text{ is non-decreasing,} \tag{96}$$

$$\int_{+0} \rho^{-2} |\log \rho|^{h_0-1} \phi(\rho)\, d\rho < \infty. \tag{97}$$

(In order to verify that (91) is a consequence of (97), introduce the new integration variable t in (97) defined by $\rho = \text{const.}\, e^{\mu t} t^{j_0}$ and note that $\rho^{-1} d\rho \sim \mu dt$ and that $\log \rho \sim \mu t$ as $t \to \infty$ (and/or $\rho \to 0$).)

Assertion (†) has the following partial converse:

(††) *Let $\mu = \mu^m$ (< 0), h_* and h_0 be defined as in* (†). *Let $F(t,y)$ be defined and continuous for $0 \leqq t < \infty$, $|y| \leqq$ const. Let there exist a continuous function $f(t)$ for large t satisfying* (94) *or let there exist a continuous function $\phi(t,\rho)$ for large t and small $\rho \geqq 0$ satisfying* (89), (90), (91) *and such that*

$$\phi(t,\rho) \text{ is non-increasing in } t \text{ (for fixed } \rho). \tag{90 bis}$$

Let $y = y(t)$ be a solution of (1) *for large t satisfying* (11) *and* (12). *Then $y = y(t)$ satisfies the asymptotic formula* (86) (*that is, the asymptotic formula* (87)-(88) *for some j_0, where $0 \leqq j_0 < h_*$, and some constants c^{qk}, not all* 0).

Under conditions (90) and (90 bis), the assumption (91) is implied by

$$\int_{+0} \rho^{-2} |\log \rho|^{h_0-1} \phi(\nu |\log \rho|, \rho)\, d\rho < \infty \text{ for some } \nu < 1/|\mu|. \tag{91 bis}$$

The proof of (††) will consist essentially of two steps: first, of showing that the conditions on $F(t,y)$ and on $y(t)$ imply

$$\int^{\infty} t^{h_0-1} \{|F(t,y(t))|/|y(t)|\}\, dt < \infty;$$

second, of showing that (11), (12) and the last relation imply the asymptotic formula (86) for $y(t)$.

The conditions of (††) on F, that is, either the condition (94) or the set of conditions (89), (90), (90 bis), (91), are sufficient for the validity of both (†) and (††). In particular, in the conservative case, where $F(t,y) \equiv F(y)$, the conditions (95)-(97) are sufficient for the truth of (†) and (††). For variants of these assertions, see Section 18 below.

It will be clear from the proofs that, under condition (94), (†) and

(††) remain valid if $\mu < 0$ is relaxed to $\mu \leqq 0$ (and $F(t,y)$ is defined for all $t \geqq 0$ and all y). The simplest case of these two theorems, where (94) is assumed, J is the zero matrix and $h_0 = 1$, is contained in the theorem of [17]. The latter allows a relaxation of (94), with $h_0 = 1$, to

$$|F(t,y)| \leqq f(t)\phi(|y|), \qquad \int^{\infty} f(t)dt < \infty, \qquad \int^{\infty} d\rho/\phi(\rho) = \infty$$

(where $f(t), \phi(\rho) > 0$ are only assumed to be continuous); for a further refinement of this case, see [9], p. 51. It is easy to see, however, that the conditions in the last formula line cannot replace (94) in (††), say, even if $J = \mu I$, where $\mu \neq 0$, I is the d by d unit matrix and $h_0 = 1$.

The binary conservative cases ($d = 2$) of (†), (††), when the real parts of the characteristic roots λ_1, λ_2 of J are equal (and negative), the elementary divisors of J are simple ($h_* = 1$) and $h_0 = 1$, are given in [18]; cf. Section 18 below for a slight refinement. The analogue of the case $\lambda_1 = \lambda_2$ (< 0) and a double elementary divisor is treated in [8], pp. 526-528 (cf. [7]), in the form

$$dy/dx = (y + x + F^2(x,y))/(x + F^1(x,y))$$

under assumptions less severe than (95)-(97), with $h_0\,(= h_*) = 2$. Perron [10] had dealt with the general equation

$$(ax + by + F^1(x,y))dy = (cx + dy + F^2(x,y))dx$$

under the assumption that F^1, F^2 satisfy Lipschitz conditions with Lipschitz constants of the order ρ^{ϵ}, for some $\epsilon > 0$, on the circle $x^2 + y^2 \leqq \rho^2$; cf. [16].

13. The case $h_0 = h_* = 1$ of (†): Preliminaries. The proof of (†) will be given first for the case $h_* = 1$ (so that the elementary divisors belonging to the eigenvalues λ_k for which $\operatorname{Re}\lambda_k = \mu\,(= \mu^m)$ are simple). In this case, the double sum (22) is a single sum over q (with $k = 1$), and so the differential inequality (25_m) can be replaced by

$$|L_{ma}' - 2\mu^m L_{ma}| \leqq d\,|F(t,y)|\,|y|. \tag{98}$$

In the proof, it can be supposed that $m > 1$. It will be clear that the case $m = 1$ is substantially simpler (since $v_1(t) \equiv 1$).

Let $y = y(t) \neq 0$ be a solution of (1) satisfying (20) and (46) for some positive $c_0\,(< \frac{1}{2})$ (hence (48)) on some interval. Let K be a positive constant so large that

$$K^{-2}e^{2\mu t} < L_{ma} < K^2 e^{2\mu t} \tag{99}$$

holds on this t-interval. Then, by (89) and (48),

$$|F(t,y(t))| \leqq \phi(t,|y(t)|) \leqq \phi(t,ce^{\mu t}), \text{ where } c^2 = 2a^{-2d}c_0^{-1}K^2. \tag{100}$$

By (48) and the first inequality in (99),

$$|y|/L_{ma} \leqq c/L_{ma}^{\frac{1}{2}} \leqq cKe^{-\mu t}, \tag{101}$$

and so (98), (100), (101) imply

$$|L_{ma}' - 2\mu^m L_{ma}| \leqq cdKe^{-\mu t}\phi(t,ce^{\mu t})L_{ma}. \tag{102}$$

If the interval $[t,t_0]$ is contained in the interval under consideration, a quadrature gives

$$|\log(e^{-2\mu t}L_{ma}(t)/e^{-2\mu t_0}L_{ma}(t_0))| \leqq cdKI(c,t), \tag{103}$$

where

$$I(c,t) = \int_t^\infty e^{-\mu u}\phi(u,ce^{\mu u})\,du. \tag{104}$$

If a, c_0 and K (hence c) are fixed, it follows from (91), where $h_0 - 1 = j_0 = 0$, that there exists a $T = T(a,c_0,K)$ such that (99) implies

$$\tfrac{1}{2} \leqq e^{-2\mu t}L_{ma}(t)/e^{-2\mu t_0}L_{ma}(t_0) \leqq 2, \text{ if } T \leqq t \leqq t_0.$$

This argument shows that if (20), (46) and (99) hold on an interval $t^0 < t \leqq t_0$, where $t^0 \geqq T$, and if

$$2K^{-2}e^{2\mu t_0} < L_{ma}(t_0) < \tfrac{1}{2}K^2e^{2\mu t_0}, \tag{105}$$

then (99) holds at $t = t^0$.

If, in addition, T is sufficiently large and $w(t_0) = 0$, then (46) holds at $t = t^0$. For, by (48), (99), (100) and $|y|^2 \geqq L_{ma}$,

$$|F(t,y(t))| \leqq |y(t)|\,Ke^{-\mu t}\phi(t,ce^{\mu t}).$$

Since $f(t) = f_K(t) = Ke^{-\mu t}\phi(t,ce^{\mu t})$ satisfies (91), hence (70), the statement concerning (46) at $t = t^0$ is a consequence of the Riccati differential inequality (37) and of the Remark following the Lemma of Section 3.

Consequently, if $w(t) = 0$ and (20), (105) hold at $t = t_0$, then, when t decreases from t_0, either the solution $y(t)$ exists on $T \leqq t \leqq t_0$ and satisfies (20), (46) and (99) on this interval or a point $t = t^0$ $(> T)$ is reached with the properties that (46) and (99) hold on $t^0 \leqq t \leqq t_0$, while (20) holds on $t^0 < t \leqq t_0$ but not at $t = t^0$.

14. The case $h_0 = h_* = 1$ of (†): Completion of the proof. A repetition of the arguments of Sections 7, 8 and 10 shows that *if $T = T(a, c_0, K)$ is sufficiently large, if $t_0 > T$, and if the numbers $y^{qk}(t_0) = y_0{}^{qk}$ satisfy* (105), *then there exist at least one set of $\Sigma h(p)$ numbers $y_0{}^{pk}$ and a solution $y = y(t)$ of* (1) *on $T \leqq t \leqq t_0$ satisfying*

$$(106_1) \qquad y^{pk}(t_0) = y_0{}^{pk}$$

and

$$(106_2) \qquad y^{qk}(t_0) = y_0{}^{qk}, \qquad (106_3) \qquad y^{rk}(t_0) = 0,$$

finally (20), (46) *and* (99) *on $T \leqq t \leqq t_0$.*

Note that (106_3) implies $w(t_0) = 0$. The fact that the last italicized statement concerns $t \leqq t_0$, while that in Section 7 concerns $t \geqq t_0$, means that the rôles of p, r, of M_m, N_{m+1} and of v, w are interchanged. The "cone" $v \geqq c_0$, that is, $N_m \geqq c_0(N_m + M_m)$ or

$$c_0 \sum_q \sum_k a^{2k} \mid y^{pk} \mid^2 \leqq (1 - c_0)\left(\sum_q \sum_k a^{2k} \mid y^{qk} \mid^2 + \sum_r \sum_k a^{2k} \mid y^{rk} \mid^2\right),$$

plays the rôle of (58) in the proof of the desired statement.

Let $y = y(t)$ be a solution of (1) on $T \leqq t \leqq t_0$ supplied by the last italicized statement. Since $h_* = 1$, the index k is 1 in any pair of indices (qk) and the corresponding differential equation for y^{q1} is

$$y^{q1\prime} = \lambda(q) y^{q1} + F^{q1}(t, y).$$

This gives $\mid (y^{q1} e^{-\lambda(q)t})' \mid \leqq e^{-\mu t} \mid F(t, y) \mid \leqq e^{-\mu t} \phi(t, \mid y(t) \mid)$, and so, in view of (48) and (99), $\mid (y^{q1} e^{-\lambda(q)t})' \mid \leqq e^{-\mu t} \phi(t, ce^{\mu t})$. Hence a quadrature over $[t, t_0]$, where $T \leqq t \leqq t_0$, gives

$$(107) \qquad \mid y^{q1}(t) e^{-\lambda(q)t} - y_0{}^{q1} e^{-\lambda(q)t_0} \mid \leqq I(c, t); \qquad \text{cf. (104).}$$

The assertion of (†), where $h_0 = 1$, is to the effect that if $\sum_q h(q) = \sum_q 1$ numbers c^{q1}, not all 0, are given, then (1) has a solution satisfying, as $t \to \infty$,

$$(108_1) \quad y^{q1}(t) = c^{q1} e^{\lambda(q)t} + o(e^{\mu t}); \qquad (108_2) \quad y^{jk}(t) = o(e^{\mu t}) \text{ if } j \neq q.$$

In terms of the given numbers c^{q1}, put

$$(109) \qquad y_0{}^{q1} = c^{q1} e^{\lambda(q)t_0}.$$

When the numbers c^{q1} are fixed, it is possible to choose a fixed, sufficiently large, constant K so that (106_2) satisfies (105).

If $t_0 = n$ is a sufficiently large integer, then, according to the last section, (1) has a solution $y = y_n(t)$ on $T \leqq t \leqq n$, where $T = T(a, c_0, K)$, satisfying (106_2), (106_3) at $t = t_0 (= n)$, and (20), (46) and (99) on $T \leqq t \leqq n$.

The inequality (107), where $T \leqq t \leqq t_0\,(=n)$, is applicable to the solution $y = y_n(t)$. Hence, $|\, y_n{}^{q1}(t)e^{-\lambda(q)t} - c^{q1}\,| \leqq I(c,t)$. The numbers $|\, y_n(T)\,|$ form therefore a bounded sequence as $n \to \infty$. Hence it is possible to select a subsequence (which can then be supposed to be the full sequence) from $y_1(t), y_2(t), \cdots$ so that

$$y(t) = \lim_{n\to\infty} y_n(t) \tag{110}$$

holds uniformly on every bounded t-interval $T \leqq t \leqq \text{const.}$ The limit $y = y(t)$ is a solution of (1) on $T \leqq t < \infty$.

Consequently, for $T \leqq t < \infty$,

$$|\, y^{q1}(t)e^{-\lambda(q)t} - c^{q1}\,| \leqq I(c,t). \tag{111}$$

The definition (104) of $I(c,t)$ and the assumption (91) of (†) show that

$$I(c,t) = o(1) \text{ as } t \to \infty, \tag{112}$$

if c is fixed. The relations (108_1) follow from (111) and (112).

The auxiliary functions $v(t)$, $w(t)$ belonging to $y(t)$ satisfy $v(t) \geqq c_0$, $w(t) \leqq \frac{1}{2}c_0$ for $T \leqq t < \infty$, since the corresponding functions of $y_n(t)$ satisfy (20), (46) for $T \leqq t \leqq t_0\,(=n)$. Hence (11) and (12) hold; cf. Section 5. The relations (88_2) follow from (11), (12) and (108_1). This completes the proof of (†) for the case $h_0\,(=h_*) = 1$.

Remark. For reference below, note that, according to the proof just completed, if $F(t,y)$ is defined for $t \geqq T$, $|\, y\,| \leqq 1 < K(t)$, where $K(T) \to \infty$ as $T \to \infty$, then (†) remains true in the case $h_0 = 1$ when $\mu = 0$ and (89)-(91) are assumed.

15. The case $h_0 > 1$: Change of variables. The proof of (†) in the cases $h_0 > 1$ will depend essentially on the case $h_0 = 1$ and a change of variables similar to that used in the proof of (**) in [6]. Let the independent variable t be replaced by $s = \log t$ (for large t and/or large s),

$$t = e^s \qquad (ds = dt/t). \tag{113}$$

Let the dependent variables be subject to a linear transformation, depending on $t = t(s)$,

$$y = Qz, \text{ where } Q = D_0D_1CD_2, \tag{114}$$

and D_0, D_1, C, D_2 are d by d matrices to be given below.

The transformation $y = D_0u$ is given by

$$y^{qk} = e^{\lambda(q)t}u^{qk} \text{ and } y^{jk} = e^{\mu t}u^{jk} \text{ if } j \neq q. \tag{115}$$

The triangular matrix D_1 depends on the fixed number β, where

$$\beta = h_0 - j_0, \tag{116}$$

and is the matrix of the linear transformation given by

$$u^{qk} = t^{k-\beta} v^{qk} \text{ and } u^{jk} = t^{1-\beta} v^{jk} \text{ if } j \neq q. \tag{117}$$

The matrix C is constant and is the matrix of the linear transformation $v = Cw$, where

$$v^{qk} = \sum_{n=1}^{k} w^{qn}/(k-n)! \text{ and } v^{jk} = w^{jk} \text{ if } j \neq q. \tag{118}$$

The change of independent variables (113) and the successive changes (115), (117), (118) of dependent variables transform the system (4) into

$$\dot{w}^{qk} = (\beta - k) w^{qk} + E^{qk}(s, w) \tag{119}$$

and, if $j \neq q$,

$$\begin{aligned} \dot{w}^{j1} &= ((\lambda(j) - \mu)t + \beta - 1) w^{j1} + E^{j1}(s, w), \\ \dot{w}^{jk} &= ((\lambda(j) - \mu)t + \beta - 1) w^{jk} + t w^{j\,k-1} + E^{jk}(s, w), \qquad k > 1, \end{aligned} \tag{120}$$

where a dot denotes differentiation with respect to s and $E(s, w)$ is the vector $t(D_0 D_1 C)^{-1} F(t, D_0 D_1 C w)$, where $t = e^s$. Cf. [6], Sections 12 and 13.

The choice of the transformation $w = D_2 z$ depends on the given j_0 determining $l(q)$ and $k(q)$. In terms of β and the numbers $l(q)$, $k(q)$, define the linear transformation $w = D_2 z$ by

$$\begin{aligned} w^{qk} &= t^{\beta-k} z^{qk} \quad \text{if} \quad l(q) \leqq k \leqq k(q), \\ w^{qk} &= z^{qk} \quad \text{if} \quad k < l(q) \text{ or } k > k(q), \\ w^{jk} &= z^{jk} \quad \text{if} \quad j \neq q. \end{aligned} \tag{121}$$

Since $t = e^s$, (101) transforms the system (119)-(120) into

$$\begin{aligned} \dot{z}^{qk} &= 0 \cdot z^{qk} + B^{qk}(s, z) \quad \text{if} \quad l(q) \leqq k \leqq k(q), \\ \dot{z}^{qk} &= (\beta - k) z^{qk} + B^{qk}(s, z) \quad \text{if} \quad k < l(q) \text{ or } k > k(q); \end{aligned} \tag{122}$$

and, if $j \neq q$,

$$\begin{aligned} \dot{z}^{j1} &= ((\lambda(j) - \mu)t + \beta - 1) z^{j1} + B^{j1}(s, z), \\ \dot{z}^{jk} &= ((\lambda(j) - \mu)t + \beta - 1) z^{jk} + t z^{j\,k-1} + B^{jk}(s, z) \qquad k > 1, \end{aligned} \tag{123}$$

where the vector $B(s,z)$ is

$$B(s,z) = tQ^{-1}F(t,Qz), \qquad t = e^s \tag{124}$$

and Q is the matrix of the transformation $y = Qz$; cf. (114).

By (115), (117), (118) and (121), the transformation $y = Qz$ is given by

$$\begin{aligned} y^{qk} &= e^{\lambda(q)t}\{t^{k-\beta}\,\Sigma_{\mathrm{I}}\, z^{qi}/(k-i)\,! + \Sigma_{\mathrm{II}}\, t^{k-i}z^{qi}/(k-i)\,!\}, \\ y^{jk} &= e^{\mu t}t^{1-\beta}z^{jk} \quad \text{if} \quad j \neq q, \end{aligned} \tag{125}$$

where Σ_{I} indicates the sum over the indices i satisfying $i \leqq k$ and either $i < l(q)$ or $i > k(q)$, and Σ_{II} is the sum over the indices i satisfying $i \leqq k$ and $l(q) \leqq i \leqq k(q)$. The transformation $z = Q^{-1}y$ is given by

$$\begin{aligned} z^{qk} &= e^{-\lambda(q)t}\sum_{i=1}^{k}(-1)^i t^{k-i}y^{qi}/(k-i)\,! \text{ if } l(q) \leqq k \leqq k(q), \\ z^{qk} &= e^{-\lambda(q)t}\sum_{i=1}^{k}(-1)^i t^{\beta-i}y^{qi}/(k-i)\,! \text{ if } k < l(q) \text{ or } k > k(q), \\ z^{qk} &= e^{-\mu t}t^{\beta-1}y^{jk} \text{ if } j \neq q. \end{aligned} \tag{126}$$

If $|\,Q\,|$ denotes the norm of Q (that is, $\max|\,Qz\,|$ for $|\,z\,| = 1$), then (125) shows that $|\,Q\,|$ is $O(e^{\mu t})$ times t^{γ}, where $\gamma = \max(h_* - \beta, h(q) - l(q))$ and the max refers to the set of q's for which $c^{qk} \neq 0$ for some k. Since $h_* - \beta = j_0 + h_* - h_0 \leqq j_0$, it follows from the last part of (88) that

$$|\,Q\,| = O(e^{\mu t}t^{j_0}) \quad \text{as} \quad t \to \infty. \tag{127}$$

By (126), $|\,Q^{-1}\,|$ is $O(e^{-\mu t})$ times t^{γ}, where $\gamma = \max(\beta - 1, k(q) - 1)$. The first part of (88) shows that $k(q) \leqq h_0 - j_0\ (= \beta)$. Hence

$$|\,Q^{-1}\,| = O(e^{-\mu t}t^{h_0 - j_0 - 1}) \quad \text{as} \quad t \to \infty. \tag{128}$$

In view of (89) and (124),

$$|\,B(s,z)\,| \leqq t\,|\,Q^{-1}\,|\,\phi(t,|\,Qz\,|). \tag{129}$$

16. Completion of the proof of (†). The relations (127)-(129) give

$$|\,B(s,z)\,| \leqq \psi(s,|\,z\,|), \tag{130}$$

where, for a sufficiently large Const.,

$$\psi(s,\rho) = \text{Const.}\, te^{-\mu t}t^{h_0 - j_0 - 1}\phi(t, \text{Const.}\, e^{\mu t}t^{j_0}\rho). \tag{131}$$

The function (131) satisfies the condition that

(132) $\psi(s,\rho)$ is non-decreasing in ρ (for fixed s)

and, by (91) and $ds = dt/t$,

$$\int^{\infty} \psi(s, \text{const.})\, ds < \infty \quad \text{for every const.} > 0. \tag{133}$$

Conditions (130), (132), (133) are the analogues of (89), (90), (91) for the case $\mu = 0$ and $h_0 - 1 = j_0 = 0$.

The fact that the coefficient of z^{qk} (in the linear terms of the first equation of (122)) is 0, rather than a negative constant, does not prevent the applicability of the case $h_0 = 1$ of (†), under assumptions (89)-(91), to the system (122)-(123); cf. the Remark at the end of Section 14. Nor does the fact that the coefficients of the linear terms on the right of the equations in (123) are not constants prevent the applicability of this assertion, since $\mathrm{Re}(\lambda(j) - \mu) \neq 0$; cf. the Remark at the end of Section 10.

Hence (130), (132) and (133) imply that, for given numbers c^{qk}, not all 0, where $l(q) \leqq k \leqq k(q)$, there is at least one solution $z = z(s)$ of (122)-(123) satisfying, as $s \to \infty$,

$$\begin{aligned} &z^{qk}(s) \to c^{qk} \text{ if } l(q) \leqq k \leqq k(q), \\ &z^{qk}(s) \to 0 \text{ if } k < l(q) \text{ or } k > k(q), \\ &z^{jk}(s) \to 0 \text{ if } j \neq q. \end{aligned} \tag{134}$$

In view of (125), the corresponding solution $y = y(t)$ of (1) satisfies (87). This proves (†).

17. Proof of (††). Let $y = y(t) \not\equiv 0$ be a solution of (1) for large t. It is clear that, under the assumption (94) on $F(t,y)$, the (vector) function $y = y(t)$ does not vanish and satisfies

$$\int^{\infty} t^{h_0 - 1} \{ |F(t, y(t))| / |y(t)| \}\, dt < \infty. \tag{135}$$

It will be verified that the same is true if $y = y(t)$ satisfies (11) and (12), where $\mu < 0$, and $F(t,y)$ satisfies (89), (90), (90 bis) and (91).

That $y(t)$ does not vanish (for large t) follows from (12). In view of (11), the auxiliary functions $v(t)$, $w(t)$ satisfy the respective limit relations (44), (60). In particular, (20) and (46) hold, where $0 < c_0 < \frac{1}{2}$, and so (48) and (49) hold for large t. Let

$$\rho(t) = L_{m a}^{\frac{1}{2}}(t) \ (> 0), \tag{136}$$

so that, by (11), there exist positive constants c, K such that

$$K\rho(t) \leqq |y(t)| \leqq c\rho(t), \tag{137}$$

and, by (12),

$$\log \rho(t) \sim \mu t \quad \text{as} \quad t \to \infty\,; \tag{138}$$

in particular,

$$t > \nu |\log c\rho(t)| \tag{139}$$

if t is sufficiently large and $\nu < 1/|\mu|$.

In view of (89) and (90)-(90 bis), it follows that

$$|F(t, y(t))|/|y(t)| \leqq \phi(\nu |\log c\rho(t)|, c\rho(t))/K\rho(t). \tag{140}$$

In order to prove (133), it will first be shown that

$$|F(t, y(t))|/|y(t)| \to 0 \quad \text{as} \quad t \to \infty. \tag{141}$$

To this end, it is sufficient to show that $\phi(\nu |\log \rho|, \rho)/\rho \to 0$ as $\rho \to 0$; actually, even more is true, namely,

$$|\log \rho|^{h_0-1} \phi(\nu |\log \rho|, \rho)/\rho \to 0 \quad \text{as} \quad \rho \to 0. \tag{142}$$

In order to see this, note that an integration by parts shows that the integral

$$\int_\rho \sigma^{-2} |\log \sigma|^{h_0-1} \psi(\sigma)\, d\sigma,$$

where $\psi(\rho) = \phi(\nu |\log \rho|, \rho)$, is

$$\psi(\rho) \int_\rho \sigma^{-2} |\log \sigma|^{h_0-1}\, d\sigma + \int_\rho \Big(\int_\sigma \tau^{-2} |\log \tau|^{h_0-1}\, d\tau \Big) d\psi(\sigma),$$

Since (90)-(90 bis) implies that $d\psi(\sigma) \geqq 0$, it follows that

$$\int_\rho \sigma^{-2} |\log \sigma|^{h_0-1} \psi(\sigma)\, d\sigma \geqq \psi(\rho) \int_\rho \sigma^{-2} |\log \sigma|^{h_0-1}\, d\sigma.$$

The factor of $\psi(\rho)$ on the right is $O(\rho^{-1} |\log \rho|^{h_0-1})$ as $\rho \to 0$. Thus (142) follows from (91 bis) (which, in turn, is a consequence of (91)).

It is seen from (141) that

$$d\rho/dt < 0 \text{ for large } t\,; \tag{143}$$

in fact, for large t,

$$(\mu - 2a)\rho < d\rho/dt < (\mu + 2a)\rho, \tag{144}$$

by (49) and (141). The inequality (143) assures that it is possible to introduce the new integration variable $\rho = \rho(t)$ into

$$\int^{\infty} \rho^{-1}(t)\, t^{h_0-1} \phi(\nu\,|\log c\rho(t)|, c\rho(t))\, dt. \tag{145}$$

In view of (91 bis), (138) and (144), the integral (145) is finite. Consequently, (135) follows from (140).

In terms of the vectors $F = (F^1, \cdots, F^d)$ and $y(t) = (y^1(t), \cdots, y^d(t))$, define, for large t, a d by d matrix $G(t) = \| g_{jk}(t) \|$ as follows:

$$g_{jk}(t) = F^j(t, y(t))\, \bar{y}^k(t) / |\, y(t)\,|^2, \tag{146}$$

where the bar denotes complex conjugation. The function (146) is continuous for large t and, by (135), satisfies condition (84). Since $y = y(t)$ is a solution of (1), it satisfies the linear system

$$y' = (J + G(t))\, y. \tag{147}$$

Hence, the assertion (86) of (††) follows from the theorem (**) of [6] on the linear systems described in Section 11 (or, equivalently, from (†) above, applied to (147), and from the superposition principle). This proves (††).

18. Variants of (†) and (††). The variant of (†) to be considered implies that in the conservative case, $F(t, y) \equiv F(y)$, of (†) the condition (96) can be relaxed to

$$|\, F(y)\,| / |\, y\,| \to 0 \quad \text{as} \quad |\, y\,| \to 0 \tag{148}$$

(whereas (95)-(97) imply the stronger condition

$$(\log |\, y\,|)^{h_0-1}\, |\, F(y)\,| / |\, y\,| \to 0 \quad \text{as} \quad |\, y\,| \to 0; \tag{149}$$

cf. the deduction of (142)). It can be remarked, however, that a disadvantage of the requirement of the existence of a $\phi(\rho)$ satisfying the set of conditions (95), (148) and (97), rather than the set (95), (96) and (97), is that the first set is not invariant under affine transformations of the variables y.

(† bis) *Assertion* (†) *remains valid if the conditions* (90), (91) *on* $\phi(t, \rho)$ *are replaced by the conditions* (7) *on* $F(t, y)$ *and the conditions* (90 bis), (91 bis) *on* $\phi(t, \rho)$.

It will be clear from the proof that (7) can be relaxed to the requirement that

$$|\, F(t, y)\,| / |\, y\,| \to 0 \quad \text{as} \quad t \to \infty \tag{150}$$

uniformly on every (t, y)-set

$$K^{-1}e^{\mu t}t^{j_0} \leqq |\, y \,| \leqq Ke^{\mu t}t^{j_0}, \tag{151}$$

where $K > 0$ is an arbitrary constant.

In view of († bis), there arises the question whether or not conditions (90), (91) in (††) can be replaced by conditions (7), (90 bis), (91 bis). This question will be answered here only in certain cases.

(†† bis) *Suppose that either* $h_* = 1$ *or* $\mu < -1$. *Then* (††) *remains valid when condition* (90) *is omitted and condition* (91) *is replaced by* (7) *and* (91 bis).

The number 1 in the inequality $\mu < -1$ is the number occurring on the subdiagonal of J.

19. Proof of († bis). Consider the change of variables (113), (114) in (1), leading to the system (122)-(123). For any solution $z = z(s)$ of (122)-(123), let $y = y(t)$ denote the solution of (1) given by $t = e^s$ and $y = Qz$.

It will first be shown that, *under condition* (7), *when* a *and* $1/K$ *are fixed* (*sufficiently small*) *positive numbers, there exist positive numbers* $T = T(a, K)$ *and* $\epsilon = \epsilon(a)$ *with the property that if* $z = z(s)$ *is a solution of* (122)-(123) *on some* s*-interval, where* $t = e^s \geqq T$, *satisfying*

$$|\, z^{qk} - c^{qk} \,| < \epsilon \text{ if } l(q) \leqq k \leqq k(q), \tag{152}$$

$$|\, z^{jk} \,| < \epsilon \text{ otherwise}, \tag{153}$$

then the corresponding solution $y = y(t)$ *of* (1) *satisfies*

$$K^{-1}e^{\mu t}t^{j_0} < |\, y \,| < Ke^{\mu t}t^{j_0}, \tag{154}$$

$$|\, |\, y \,|' - \mu\, |\, y \,|\, | \leqq a\, |\, y \,|. \tag{155}$$

In (154), j_0 denotes the maximum of the differences $h(q) - k$ for which $c^{qk} \neq 0$; so that it can be supposed that $j_0 = h(q) - l(q)$ for some q. Let $q = q_0$ denote any q satisfying $j_0 = h(q) - l(q)$. Let

$$L^0 = \Sigma^0\, |\, y^{qh} \,|^2, \tag{156}$$

where Σ^0 denotes a sum over those $q = q_0$. Let K be so large that (152), (153) imply

$$(\tfrac{1}{2}K)^{-1} < \Sigma^0\, |\, z^{qk} \,|^2 \leqq |\, z \,|^2 < \tfrac{1}{2}K. \tag{157}$$

Then (125), where $h_* - \beta \leqq j_0$, and (152) show that

$$|\, y^{jk} \,| \leqq \epsilon e^{\mu t}t^{j_0} \text{ if } j \neq q,$$

$$|\, y^{qk} \,| \leqq \tfrac{1}{2}Kde^{\mu t}t^{j_0 - 1} \text{ if } q = q_0,\ k < h \text{ or if } q \neq q_0,\ k \leqq h,$$

and that $|y^{qh}(t)|$, when $q=q_0$, differs from $|c^{qh}|e^{\mu t}t^{j_0}/(h-l)!$ by terms which are majorized either by $\epsilon e^{\mu t}t^{j_0}$ or by $Ke^{\mu t}t^{j_0-1}$. Hence, if $\epsilon=\epsilon(a)$ is sufficiently small, $T=T(a,K)$ sufficiently large and $t\geqq T$, then (154) holds and

$$L^0(t)\leqq|y(t)|^2\leqq(1+a^2)L^0(t). \tag{158}$$

From the differential system (4) and from $(|y^{jk}|^2)'=\bar{y}^{jk}y^{jk'}+\bar{y}^{jk'}y^{jk}$, it follows that $||y|^{2'}-2\mu L_0(t)|$ is majorized by a constant multiple of the sum of $|y(t)|^2-L^0(t)$, $|y(t)|(|y(t)|^2-L^0(t))^{\frac{1}{2}}$ and $F(t,y(t))|\cdot|y(t)|$. In view of (7) and (158), the inequality (155) follows if a is sufficiently small and $T(\geqq t)$ is sufficiently large. This proves the italicized statement concerning (152)-(155).

The proof of († bis) can now be completed along the lines used in Sections 13-14 in the proof of (†) when $h_0=1$. To this end, define L_a, $M=M_{ma}$, $N=N_{m+1\,a}$ by the formulae

$$L_a=a^2\sum_q\sum_{k=l(q)}^{k(q)}|z^{qk}|^2, \tag{159}$$

$$M=\sum_p\sum_k a^{2k}|z^{pk}|^2+a^2\sum_q\sum_{k>k(q)}|z^{qk}|^2, \tag{160}$$

$$N=\sum_r\sum_k a^{2k}|z^{pk}|^2+a^2\sum_q\sum_{k<l(q)}|z^{qk}|^2. \tag{161}$$

Thus, the analogues of the auxiliary functions $v=v(s)$, $w=w(s)$ are

$$v=(N+L_a)/(M+L_a+N),\qquad w=N/(M+L_a+N). \tag{162}$$

These satisfy the Riccati differential inequality

$$\dot{u}\geqq\tfrac{1}{2}u(1-u)-d^2a^{-2d}B(s,z(s))|/|z(s)|,\qquad (u=v\text{ or }u=w); \tag{163}$$

cf. the derivation of (30), (33).

Consider a solution $z=z(s)$ of (122)-(123) satisfying the (possibly incomplete) set of initial conditions

$$z^{qk}(s_0)=c^{qk}\text{ if }l(q)\leqq k\leqq k(q);\qquad N(s_0)=0. \tag{164}$$

Suppose further that (152) and that (20) and (46) hold on some s-interval (s,s_0) for some c_0, where $0<c_0<\frac{1}{2}$. Then the analogue of (47) gives

$$M+N<(2c_0^{-1}-1)L_a \tag{165}$$

on this interval. Thus, if $c_0=c_0(a)$ is sufficiently near $\frac{1}{2}$, the inequalities (152) on $(s,s_0]$ imply (153) there.

A quadrature of the equations (122) shows that

$$(166) \qquad |z^{qk}(s) - c^{qk}| \leqq \int_s^{s_0} |B(s, t(s))|\, ds, \qquad l(q) \leqq k \leqq k(q).$$

In view of (128), (129) and $ds = dt/t$,

$$\int_s^{s_0} |B(s, z(s))|\, ds \leqq \text{Const.} \int_t^{t_0} e^{-\mu t} t^{h_0 - j_0 - 1}\, \phi(t, |y(t)|)\, dt,$$

where $t_0 = e^{s_0}$, $t = e^s$.

By virtue of the last italicized statement, (154)-(155) hold on the interval under consideration. If a is so small that $\mu + a < 0$, then, by (155), $|y(t)|$ is decreasing on the interval under consideration, and so it is possible to introduce $\rho = |y(t)|$ as a new integration variable in the last integral. By (155), $dt \leqq |\mu + a|^{-1} |d\rho|/\rho$ and, by (154), $e^{-\mu t} t^{-j_0} < K/\rho$. In addition, $t > \nu |\log \rho|$ for $\nu = 1/|\mu - a|$ and $t \geqq T$ if $T = T(a, K)$ is sufficiently large. Hence

$$(167) \qquad \int_s^{s_0} |B(s, z(s))|\, ds \leqq cI(Ke^{\mu t} t^{j_0}),$$

where $c = c(a, K)$ is a constant and

$$(168) \qquad I(\sigma) = \int_{+0}^{\sigma} \rho^{-2} |\log \rho|^{h_0 - 1}\, \phi(\nu |\log \rho|, \rho)\, d\rho.$$

It follows from (166), (167) and (168), that if (152), (20) and (46) hold on an interval $s^0 < s \leqq s_0$ (that is, $t^0 < t \leqq t_0$) and if t^0 exceeds a sufficiently large $T = T(a, k)$, then (152) also holds at $s = s^0$ (that is, at $t = t^0$). In view of the case $u = w$ of (163), where $|z| > 1/K$ (cf. (157)) on $t^0 < t \leqq t_0$, and of (167), (168), the Lemma of Section 3 implies that if $T = T(a, c_0, K)$ is suitably chosen and $t^0 \geqq T$, then (46) holds at t^0, since $w(t_0) = 0$, by (164).

These facts and the estimates (166), (167) show that the proof of († bis) can now be completed by the arguments of Section 14 used to complete the proof of the case $h_0 = 1$ of (†).

20. Proof of (†† bis). It is clear from the proof of (††) that in order to verify the variant of (††), it is sufficient to show that, under the conditions of these statements, (135) holds or, what is more, that

$$\int^{\infty} t^{h_0-1}\phi(t, |y(t)|)\,dt < \infty. \tag{169}$$

To this end, it is enough to show that, for large t,

$$|y(t)|' < -\text{const.}\,|y(t)|, \qquad (\text{const.} > 0). \tag{170}$$

For this means that it is possible to introduce the new integration variable $\rho = |y(t)|$ in the last integral and that $dt \leqq \text{const.}\,|d\rho|/\rho$. This fact, when combined with

$$\log|y(t)| \sim \mu t \quad \text{as} \quad t \to \infty \tag{171}$$

(cf. (11), (12)), and assumption (91 bis) imply (169).

Condition (11) means that, as $t \to \infty$,

$$|y(t)|^2 \sim L(t), \tag{172}$$

where $L = L_m$ is defined in (6). It is clear from the differential equations (4) that $|(|y(t)|^2)' - 2\mu L(t)|$ is majorized by

$$\text{const.}\,\{(|y(t)|^2 - L(t)) + |F(t,y)|\,|y|\} + 2\sum_q \sum_{k=2}^{h} |y^{qk}y^{q\,k-1}|.$$

If $h_* = 1$, the last term is missing. If $h_* \geqq 1$, the last term is not greater than $2L(t)$. Hence (170), for large t, follows from (172) if either $h_* = 1$ or $\mu < -1$. This proves (†† bis).

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] O. Dunkel, "Regular singular points of a system of homogeneous linear differential equations of the first order," *Proceedings of the American Academy of Arts and Sciences*, vol. 38 (1902-03), pp. 341-370.

[2] P. Hartman, "Unrestricted solution fields of almost separable differential equations," *Transactions of the American Mathematical Society*, vol. 63 (1948), pp. 560-580.

[3] ——— and A. Wintner, "On the asymptotic behavior of the solutions of a non

linear differential equation," *American Journal of Mathematics*, vol. 68 (1946), pp. 301-308.

[4] ——— and A. Wintner, "On the behavior of the solutions of real binary differential systems at singular points," *ibid.*, vol. 75 (1953), pp. 117-126.

[5] ——— and A. Wintner, "On curves defined by binary non-conservative differential systems," *ibid.*, vol. 76 (1954), pp. 497-501.

[6] ——— and A. Wintner, "Asymptotic integrations of linear differential equations," *ibid.*, vol. 77 (1955), pp. 45-87.

[7] E. Lonn, "Knoteninvarianz bei Differentialgleichungen," *Jahresbericht der Deutschen Mathematiker-Vereinigung*, vol. 43 (1934), pp. 232-237.

[8] ———, "Über singuläre Punkte gewöhnlicher Differentialgleichungen," *Mathematische Zeitschrift*, vol. 44 (1939), pp. 507-530.

[9] S. Łojasiewicz, "Sur l'allure asymptotique des intégrales du système d'équations différentielles au voisinage de point singulier," *Annales Polonici Mathematici*, vol. 1 (1954), pp. 34-72.

[10] O. Perron, "Ueber die Gestalt der Integralkurven einer Differentialgleichung erster Ordnung in der Umgebung eines singulären Punktes," *Mathematische Zeitschrift*, vol. 15 (1922), pp. 121-146 and vol. 16 (1923), pp. 273-295.

[11] ———, "Über Stabilität und asymptotisches Verhalten der Integrale von Differentialgleichungssystemen," *ibid.*, vol. 29 (1929), pp. 129-160.

[12] ———, "Die Stabilitätsfrage bei Differentialgleichungen," *ibid.*, vol. 32 (1930), pp. 703-728.

[13] Z. Szmydtówna, "Sur l'allure asymptotique des intégrales des équations différentielles ordinaires," *Annales de la Société Polonaise de Mathématique*, vol. 24 (1951), No. 2, pp. 17-34.

[14] T. Ważewski, "Sur un principe topologique de l'examen de l'allure asymptotique des intégrales des équations différentielles ordinaires," *ibid.*, vol. 20 (1947), pp. 279-313.

[15] ———, "Sur les intégrales d'un système d'équations différentielles tangentes aux hyperplans caractéristiques issus du point singulier," *ibid.*, vol. 21 (1948), pp. 277-297.

[16] A. Wintner, "Asymptotic integration constants in the singularity of Briot-Bouquet," *American Journal of Mathematics*, vol. 68 (1946), pp. 293-300.

[17] ———, "Asymptotic integrations constants," *ibid.*, vol. 68 (1946), pp. 553-559.

[18] ———, "Vortices and modes," *ibid.*, vol. 69 (1947), pp. 816-824.

PRIMITIVE ALTERNATIVE RINGS AND SEMI-SIMPLICITY.*

By Erwin Kleinfeld.

1. Introduction. Jacobson has characterized primitive, associative rings as dense rings of linear transformations on vector spaces over division rings and developed a structure theory for semi-simple associative rings without chain conditions [3], which generalizes the well known Wedderburn theorems. It seems natural enough to ask whether this theory extends to alternative rings. This question was first proposed by Smiley ([8], question 5). Kaplansky has shown [4] that $\mathfrak{p}$-regular, primitive, alternative rings are either associative or Cayley-Dickson algebras. He also conjectured that the assumption of $\mathfrak{p}$-regularity would prove to be superfluous. The present note shows that indeed this is the case.

The method of attack is to prove a primitive, alternative ring either associative or simple. This in turn is facilitated by the ease with which one can construct two-sided ideals within right ideals. As a by-product we obtain an elementary proof that an alternative, non-associative ring without proper two-sided ideals has no proper one-sided ideals. This accentuates how different the associative theory is from the alternative, for while it is an elementary exercise to show that an associative ring without proper right ideals is either trivial or a division ring, the last mentioned result does not seem to make the classification of simple, alternative rings any easier. Indeed it is still an open question whether alternative, non-associative nil rings without proper ideals exist.

The notion of primitive ring lends itself to a structure theory for semi-simple rings. Every alternative ring modulo its radical turns out to be a subdirect sum of primitive rings.

2. Definitions and identities. In order to facilitate the reading of the proofs we mention some of the essential definitions and identities. For background material as well as development of the identities we refer the reader to [2].

We define $(a, b, c) = ab \cdot c - a \cdot bc$, and $(a, b) = ab - ba$, for arbitrary elements, a, b, c of a ring R. If R is assumed to be alternative then (a, b, c) becomes a skew-symmetric function in its three variables.

The function $f(w, x, y, z)$ is defined by the equation

$$(wx, y, z) = f(w, x, y, z) + x(w, y, z) + (x, y, z)w.$$

* Received May 25, 1955.

$f(w,x,y,z)$ turns out to be skew-symmetric in its four variables. Furthermore it satisfies the equation $f(w,x,y,z) = ((w,x),y,z) + (w,x,(y,z))$.

The nucleus N of R is defined as the set of all n in R with the property $(n,R,R) = 0$. The center C of R is defined as the set of all c in N with the property $(c,R) = 0$. The connection between N and C is explained by Lemma 2.3 of [5].

An alternative ring is defined to be simple in case it has no proper two-sided ideals and is not a nil ring.

An alternative ring is defined to be primitive in case it contains a regular, maximal right ideal, which contains no two-sided ideal of the ring other than the zero ideal.

The radical of an alternative ring is defined as the intersection of its regular, maximal right ideals.

3. Construction of ideals. If A is any right ideal of an alternative ring R then we define an element s to be special relative to A in case (i) s is an element of A and (ii) $Rs \subset A$. The following Lemma justifies this definition.

LEMMA 1. *For any right ideal A of an alternative ring R, the special elements form an ideal of R.*

Proof. Let s, s' be any two special elements. It follows directly from the definition that $s - s'$ must then be special. If x and y are arbitrary elements of R, consider first xs. By definition xs is in A, while $y \cdot xs = -(y,x,s) + yx \cdot s$. Since A is a right ideal of R and s is a member of A, it follows that $(y,x,s) = (s,y,x) = sy \cdot x - s \cdot yx$ is also in A. We have shown that $y \cdot xs$ is an element of A and consequently xs is special. So far it has been established that the special elements form a left ideal of R. But also $y \cdot sx = -(y,s,x) + ys \cdot x$. Since ys is in A, so is $ys \cdot x$ and therefore $y \cdot sx$. But then the special elements form a two-sided ideal, as was to be shown.

In the next Lemma we develop an important property of the ideal of special elements.

LEMMA 2. *For any right ideal A of an alternative ring R all elements of the form (A,A,R) are special.*

Proof. As before $(A,A,R) \subset A$. Let us consider elements of the form $R(A,A,R)$. Select any a, a' in A and r, s in R. Then

$$s(r,a,a') = (rs,a,a') - f(r,s,a,a') - (s,a,a')r.$$

Clearly (rs, a, a') is in A, as well as $(s, a, a')r$. Since

$$f(r, s, a, a') = ((r, s), a, a') + (r, s, (a, a')),$$

we obtain also that $f(r, s, a, a')$ is a member of A. Consequently $R(A, A, R) \subset A$, which makes all elements of the form (A, A, R) special.

At this point we obtain the following result:

THEOREM 1. *If R is an alternative, non-associative ring which has no proper two-sided ideals, then it has no proper left or right ideals.*

Proof. Let us suppose that A is a proper right ideal of R. Then by virtue of Lemma 1 the special elements of A consist of just the zero element. Then we may use Lemma 2 to get $(A, A, R) = 0$. From Theorem 3.2 of [5] it follows that the nucleus N and the center C of R coincide. This in turn is used to prove Lemma 3.3 of [5], which asserts that if $(x, y, R) = 0$ then $(x, y) = 0$. From this we conclude that $(A, A) = 0$. In other words A must be an associative, commutative subring of R. For any elements r, s in R and a, a' in A,

$$f(r, s, a, a') = (rs, a, a') - s(r, a, a') - (s, a, a')r = 0,$$

thereby making $f(A, A, R, R) = 0$. As a consequence of this

$$0 = (ar, s, a') = f(a, r, s, a') + r(a, s, a') + (r, s, a')a,$$

so that $(r, s, a')a = 0$. This implies that $(R, R, A)A = 0$. Since $(R, R, A) \subset A$, while A is commutative, also $A(R, R, A) = 0$. But then

$$(aa', r, s) = f(a, a', r, s) + a'(a, r, s) + (a', r, s)a = 0.$$

Consequently $(A^2, R, R) = 0$ and A^2 is in the center of R. In that case $RA^2 = A^2R \subset A$, and hence all elements of the form A^2 are special. This of course implies again $A^2 = 0$. Now $A + RA$ is a two-sided ideal of R. Since it contains $A \neq 0$, it must be all of R. But then $RA = (A + RA)A = 0$. Therefore A is a two-sided ideal of R. But this contradicts the original assumption that R has no proper two-sided ideal and the proof is completed.

This result is new only to the extent that it applies to nil rings. For we have shown elsewhere [6] that alternative, non-associative, simple rings are Cayley-Dickson algebras. The present proof however is quite elementary. In fact we may use it to establish the well known result that Cayley-Dickson algebras have no proper left or right ideals. The question of whether alternative, non-associative, nil rings without proper ideals exist is an open one. For aesthetic reasons one would like to be able to answer this in the negative.

4. Primitive rings. Throughout this section R will denote an alternative, non-associative, primitive ring and A a regular, maximal right ideal of R, which contains no two-sided ideal of R other than the zero ideal. Again it is clear that the special elements of A are all zero, so that $(A, A, R) = 0$ and $(R, R, A)A = 0$, using the same arguments as before. Next we prove

LEMMA 3. *If an element t of R has the property $Rt = 0$, then $t = 0$.*

Proof. The regularity of A assures the existence of an element g in R with the property that for every x in R $gx - x$ is in A. Then in particular $gt - t$ is in A. But $gt = 0$, so that t itself must be in A. In that case t is special relative to A and $t = 0$.

The next step is to show that R is prime.

LEMMA 4. *If B and B' are ideals of R such that $BB' = 0$, then either $B = 0$ or $B' = 0$.*

Proof. Suppose that $B \neq 0$. Then B cannot be contained in A. Since A is a maximal, right ideal and $A + B$ is a right ideal properly containing A, we must have $A + B = R$. Then $RB' = (A + B)B' = AB' + BB' = AB' \subset A$. But clearly RB' is an ideal of R, so that $RB' = 0$. Lemma 3 now tells us that $B' = 0$, which is what we wanted to show.

Denote by N the nucleus of R and by C the center of R. Then with the aid of Lemma 2.3 of [5] we prove the following results.

LEMMA 5. (i) $N = C$. (ii) *if a, b are elements of R such that $(a, b, R) = 0$, then $(a, b) = 0$.*

Proof. Suppose S is any non-zero subset of R and suppose the two-sided annihilators of S form an ideal T of R. Then T must be the zero ideal. For since T is an ideal, the two-sided annihilators of it also form an ideal, which moreover contains S. Then by virtue of Lemma 4, T must be zero.

If we assume $N \neq C$ then $(N, R) \neq 0$. Since $(N, R) \subset N$, the annihilators of (N, R) form an ideal, which must then be the zero ideal. From this point on we appeal to [5]. Let x be an arbitrary element of R. Then every element of the form $((R, R, R), R, R)$ or of the form $(x, R, R)(x, R, R)$ is in this ideal. The latter leads us to conclude that if u is any element of the form (R, R, R), then u is an element of N and $u^2 = 0$. The annihilators of (u, R) form an ideal, which contains u since $(x, R, R)(x, R, R) = 0$. If $u \neq 0$ then $(u, R) = 0$, so in either case u is an element of C. Note that Ru is an ideal of R and that $(Ru)^2 = 0$. Using Lemma 4 we obtain $Ru = 0$,

so that with the help of Lemma 3 $u = 0$. Since the argument holds for any associator u, R must be associative, which is contrary to assumption. The contradiction arose from the assumption that $N \neq C$, hence (i) is established.

Let c, c' be arbitrary elements of C. Then Rc and Rc' are ideals of R. If in addition $cc' = 0$, then $(Rc)(Rc') = 0$ and as before either $Rc = 0$ or $Rc' = 0$. Consequently either $c = 0$ or $c' = 0$. In other words C has no divisors of zero. Assume now that $(a, b, R) = 0$. Then as before it follows that $f(a, b, R, R) = 0$, and then $((a, b), R, R) = 0$, which means that (a, b) belongs to the set $N = C$. But $(a, ab, R) = 0$, so that $(a, ab) = a(a, b)$ also belongs to C. Then $0 = (a(a, b), b) = (a, b)^2$, and from the previous remark we may conclude that $(a, b) = 0$. This concludes the proof of the lemma.

Now for the main result:

THEOREM 2. *Every primitive, alternative, non-associative ring is a Cayley-Dickson algebra.*

Proof. Using $(A, A, R) = 0$, $(A, R, R)A = 0$ and Lemma 5 we obtain $(A, A) = 0$, $f(A, A, R, R) = 0$, $(A^2, R, R) = 0$ and $A^2 \subset C$. Suppose c belongs to the intersection of A and C. Then $Rc = cR \subset A$. Since Rc is an ideal of R, we get $Rc = 0$, which implies $c = 0$. Since all elements of the form A^2 belong to this intersection, $A^2 = 0$. Again either A is an ideal of R, in which case $A = 0$, or $A + RA$, which is an ideal of R, contains A properly and so $A + RA = R$. But if the latter is true then $RA = (A + RA)A = 0$, so in any case $A = 0$. As A was chosen to be a maximal right ideal of R this simply means that R has no proper right ideals. Also g now becomes a left unit of R and hence R cannot be nil. Then R is a simple ring and therefore a Cayley-Dickson algebra.

5. Semi-simple rings. Having defined the radical Q of an alternative ring R as the intersection of its regular, maximal right ideals, we shall call R semi-simple in case $Q = 0$. An ideal of R will be said to be primitive in case R modulo this ideal is a primitive ring. It is clear that the intersection of the primitive ideals is contained in Q and we proceed to prove equality using our knowledge of primitive rings. It will suffice to show that in a primitive ring the intersection of all regular, maximal right ideals is zero. We have already seen that Cayley-Dickson algebras contain no proper right ideals, while it follows from Jacobson's density theorem for associative, primitive rings that the intersection of all regular, maximal right ideals is zero. For take a fixed representation of R as a dense ring of linear transformations. Then the annihilators of any non-zero vector a form a regular,

maximal right ideal (an element sending a into a acting as left unit) and the intersection of all of these is clearly zero. We have proved

THEOREM 3. *The radical Q of any alternative ring R is a two-sided ideal having the property that R/Q is a subdirect sum of Cayley-Dickson algebras and primitive, associative rings.*

The radical Q has also been characterized by Brown [1] as the set of all elements a in R such that every element b in the ideal generated by a has the property that $(1-b)R$ generates R as a right ideal. Kaplansky has shown [4] that the radical defined by Smiley [7] is always contained in Q and coincides with Q in the case of a Zorn ring. Whether it coincides in an arbitrary alternative ring R is an open question. No general structure theory for R modulo its Smiley radical seems to be known.

In conclusion we point out some immediate applications of Theorem 3. Various associativity conditions that hold for Cayley-Dickson algebras now carry over to semi-simple alternative rings. For example it follows that the square of every associator of a semi-simple alternative ring lies in its center while the square of every commutator lies in its nucleus. For arbitrary alternative rings it is known only that the fourth power of every commutator lies in its nucleus [6].

OHIO STATE UNIVERSITY.

REFERENCES.

[1] B. Brown, "An extension of the Jacobson radical," *Proceedings of the American Mathematical Society*, vol. 2 (1951), pp. 114-117.

[2] R. H. Bruck and E. Kleinfeld, "The structure of alternative division rings," *Proceedings of the American Mathematical Society*, vol. 2 (1951), pp. 878-890.

[3] N. Jacobson, "The radical and semi-simplicity for arbitrary rings," *American Journal of Mathematics*, vol. 67 (1945), pp. 300-320.

[4] I. Kaplansky, "Semi-simple alternative rings," *Portugaliae Mathematica*, vol. 10 (1951), pp. 37-50.

[5] E. Kleinfeld, "An extension of the theorem on alternative division rings," *Proceedings of the American Mathematical Society*, vol. 3 (1952), pp. 348-351.

[6] ———, "Simple alternative rings," *Annals of Mathematics*, vol. 58 (1953), pp. 544-547.

[7] M. F. Smiley, "The radical of an alternative ring," *Annals of Mathematics*, vol. 49 (1948), pp. 702-709.

[8] ———, "Some questions concerning alternative rings," *Bulletin of the American Mathematical Society*, vol. 57 (1951), pp. 36-43.

[9] M. Zorn, "Theorie der alternativen Ringe," *Abhandlungen des Mathematischen Seminars der Hamburgischen Universität*, vol. 8 (1930), pp. 123-147.

APPENDIX TO THE PAPER "MEAN PERIODIC FUNCTIONS I".*

By LEON EHRENPREIS**

In our paper, "Mean Periodic Functions I" (see [2]) we considered the problem of whether an entire function f of n complex variables can be approximated by linear combinations of the exponential monomials which occur as the limits of its translates. The following result was proven: Let $\mathcal{H}$ denote the space of entire functions of n complex variables with its usual topology (see [3]), and let V be any variety in $\mathcal{H}$, i.e., V is a closed linear subspace of $\mathcal{H}$ with $V \neq \{0\}$ which is closed with respect to translation. Denote by $\mathcal{H}'$ the dual of $\mathcal{H}$ with the topology of uniform convergence on the bounded sets of $\mathcal{H}$; $\mathcal{H}'$ is a commutative topological ring under convolution. Call V' the set of all $S \in \mathcal{H}'$ such that $S \cdot f = 0$ for all $f \in V$. It is readily verified that V' is a closed ideal in $\mathcal{H}'$ which we call the annihilator ideal of V. Then we showed that if V' is a principal ideal (that is, V' has one generator as a closed ideal) then every $f \in V$ is the limit of the linear combinations of the exponential *monomials* in V. In the present paper we shall show that such is not the case if V' is not principal. (See [2], Problem 2, p. 306).

We showed in [2] that the question of whether or not every variety V is the closure of the linear combinations of the exponential monomials in V is equivalent to the following problem on ideals in rings of analytic functions: Denote by $\boldsymbol{H}'$ the Fourier transform of $\mathcal{H}'$, so $\boldsymbol{H}'$ is the ring of all entire functions of exponential type of n complex variables $Z_1, Z_2, \cdots, Z_n$; the topology of $\boldsymbol{H}'$ is described in [2]. Let I be a closed ideal in $\boldsymbol{H}'$; is I determined by its cospectrum (see [2])? That is, suppose $F \in \boldsymbol{H}'$ has the property that, for every point z in complex Euclidean n space and any non-negative integers $r_1, r_2, \cdots, r_n$, the conditions

$$[(\partial^{r_1'+r_2'+\cdots+r_n'}/\partial Z_1^{r_1'}\partial Z_2^{r_2'}\cdots\partial Z_n^{r_n'})H](z) = 0 \tag{1}$$

for all $H \in I$ whenever $0 \leqq r_1' \leqq r_1, 0 \leqq r_2' \leqq r_2, \cdots, 0 \leqq r_n' \leqq r_n$ imply

$$[(\partial^{r_1+r_2+\cdots+r_n}/\partial Z_1^{r_1}\partial Z_2^{r_2}\cdots\partial Z_n^{r_n})F](z) = 0; \tag{2}$$

is $F \in I$? We shall produce a closed ideal I in $\boldsymbol{H}'$ and a function $F \in \boldsymbol{H}'$ such that conditions (1) imply (2), but $F \notin I$.

* Received July 22, 1955.

** Work supported by National Science Foundation Grant NSF5-G1010.

(At this point we should like to note some error in the terminology used in [2]. We have often used the term "exponential polynomial in V" for "linear combination of the exponential monomials of V." The spectrum of a variety V as defined in [2] corresponds to the exponential monomials in V. Thus, problem 2 on p. 306 should be formulated as follows: "Is every variety determined by its exponential monomials, that is, if W is a variety and $f \in W$, is f the limit of the linear combinations of the exponential monomials of W?" This is equivalent to the statement about cospectra of ideals made there. Thus, our proof of Theorem 9 on p. 321 really proves that V is determined by its exponential *monomials*—a much stronger statement than that made in Theorem 9. The reader will have no difficulty in correcting the other cases of this error in notation.)

Indeed, let $n = \overset{2}{\aleph}_0$ and let I be the closed ideal in $\boldsymbol{H}'$ generated by the functions $H_1(z_1, z_2) = z_1 + z_2$, $H_2(z_1, z_2) = z_1 z_2$. Since the topology of H' is stronger than pointwise convergence, every $H \in I$ must vanish at the origin. On the other hand, the only common zero of H_1 and H_2 is clearly the point $(0,0)$. Moreover, $[(\partial/\partial Z_1)H_1](0,0) = 1$, $[(\partial/\partial z_2)H_1](0,0) = 1$. Thus, in order to construct our example, we need only produce one function $F \in \boldsymbol{H}'$ which vanishes at $(0,0)$ but does not belong to I. We claim that such a function is $F(z_1, z_2) = z_1$.

To see that $z_1 \notin I$, denote by I^* the (algebraic) ideal generated by H_1 and H_2 in $\mathfrak{A}$. Then by a result of H. Cartan, I^* is closed in $\mathfrak{A}$ (see [1], p. 62, Theorem 11). On the other hand, by [2], p. 297, Proposition 4, the topology of $\boldsymbol{H}'$ is stronger than that induced by $\mathfrak{A}$. Thus, $I^* \supset I$. But it is clear that there do not even exist formal power series G_1, G_2 in two variables such that $H_1 G_1 + H_2 G_2 = F$. Thus, $F \notin I^*$, so $F \notin I$. This completes the proof of our assertion.

Remark 1. Let us consider the change of coordinates $Z_1 \to X_1 - X_2$, $Z_2 \to X_2$. Then

$$(3) \qquad H_1 \to H_1 = X_1, \qquad H_2 \to H_2 = X_1 X_2 - X_2^2, \qquad F \to F = X_1 - X_2.$$

Thus, by (5), $[(\partial/\partial X_2)H_1](0) = 0$, $[(\partial/\partial X_2)H_2](0) = 0$, but $[(\partial/\partial X_2)F](0) = 1 \neq 0$. (This gives another proof that $F \notin I$.) Thus, though we have shown that the answer to problem 2 of [2], p. 306 (as amended above) is in the negative, it seems reasonable to expect that the answer to the following problem is in the affirmative:

Problem 2′. Let J be a closed ideal in $\boldsymbol{H}'$, and let $G \in \boldsymbol{H}'$. Suppose not only that conditions (1) (for J instead of I) imply (2) (for G instead

of F) but that the similar property holds after an arbitrary affine change of coordinates. Then, is $G \in J$?

An affirmative answer to Problem 2′ would clearly imply that every variety is determined by its exponential polynomials; the converse is, however, not obvious. Problem 2′ will be discussed in detail in a later publication.

Remark 2. It is a simple matter to extend the above to the case of more than two complex variables. Moreover, the space $\mathcal{H}$ can be replaced by essentially any space such that the Fourier transform of its dual is a space of entire functions. In particular, we may replace $\mathcal{H}$ by the space $\mathcal{E}$ of Schwartz (see [3], [4]).

THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] H. Cartan, "Idéaux et modules de fonctions analytiques de n variables complexes," *Bulletin de la Société Mathématique de France*, vol. 78 (1950), pp. 28-64.

[2] L. Ehrenpreis, "Mean periodic functions, Part I, Varieties whose annihilator ideal is principal," *American Journal of Mathematics*, vol. 77 (1955), pp. 293-328.

[3] L. Schwartz, "Theorie générale des fonctions moyenne-périodiques," *Annals of Mathematics*, vol. 48 (1947), pp. 857-929.

[4] ———, *Théorie des Distributions*, vol. I and II, Paris, 1951-52.

[5] ———, "Analyse et synthèse harmonique dans les espaces de distributions," *Canadian Journal of Mathematics*, vol. 3 (1951), pp. 503-512.

ON THE REGULAR SOLUTIONS AT A POINT OF SINGULARITY OF A SYSTEM OF NON-LINEAR DIFFERENTIAL EQUATIONS.*

By ROBERT W. BASS.[1]

1. Introduction. A theorem of Lettenmeyer [3] (which is a generalization of a theorem of Perron [5], reproduced in Bieberbach's recent *Theorie der gewöhnlichen Differentialgleichungen auf funktionentheoretischer Grundlage* (1953), pp. 161-172) concerns the existence of solutions of a system of linear differential equations

$$z^{s_i} w_i'(z) = \sum_{j=1}^{n} a_{ij}(z) w_j, \qquad (i = 1, \cdots, n), \tag{1}$$

where $w_i' = dw_i/dz$. In these equations s_i $(i = 1, \cdots, n)$ are non-negative integers such that

$$s = s_1 + s_2 + \cdots + s_n < n, \tag{2}$$

and the $a_{ij}(z)$ $(i, j = 1, \cdots, n)$ are regular in some neighborhood of $z = 0$. Then ([3]) the system (1) has at least $n - s$ linearly independent solutions regular at the origin.

The object of the present paper is to generalize this result of Lettenmeyer to the non-linear system

$$z^{s_i} w_i'(z) = f_i(z, w_1, \cdots, w_n), \qquad (i = 1, \cdots, n), \tag{3}$$

in such a manner as to include the classical Cauchy-Poincaré theorem as a special case. This problem was suggested to me by Professor Wintner.

To be precise, let $f_i(z, w_1, \cdots, w_n)$ be n functions of $n+1$ variables, regular on some neighborhood of $(z, w_1, \cdots, w_n) = (0, 0, \cdots, 0)$ and such that their power-series representations converge absolutely on the closed product set

$$|z| \leqq r_1, \qquad (3a) \qquad |w_i| \leqq r_2, \qquad (3b) \qquad (i = 1, \cdots, n).$$

Suppose further that $f_i(z, 0, 0, \cdots, 0) \equiv 0$, so that the f_i can be written in the form

$$f_i = \sum_{j=1}^{n} a_{ij}(z) w_j + g_i(z, w_1, \cdots, w_n), \qquad (i = 1, \cdots, n), \tag{4}$$

* Received May 13, 1955.

[1] National Science Foundation Predoctoral Fellow.

where the g_i are power series which converge absolutely on (3a), (3b) and contain only terms of the second degree or higher in the w_i.

THEOREM 1. *Let* (2) *and* (4) *hold. Then the system of differential equations* (3) *admits of at least an* $(n-s)$*-parameter algebroid family of solutions regular at* $z=0$. *The solutions will have algebroid dependence on any additional parameters in which the* f_i *are given regular.* (The precise meaning of "k-parameter algebroid family" and "algebroid dependence" will be set forth in Section 2.)

The technique of the proof constitutes an application of Wintner's theory ([7]) of analytic functions of infinitely many variables and of infinite implicit systems. We first suppose that (3) has a regular solution of the form

$$w_i(z) = \sum_{j=0}^{\infty} d_{ij} z^j, \qquad (i = 1, \cdots, n). \tag{5}$$

Then insert (5) in (3) and compare coefficients to obtain the necessary conditions

$$(\nu - s_i + 1)\, d_{i\,\nu-s_i+1} = f_{i\nu}\,(d_{j0}, d_{j1}, \cdots, d_{j\nu}; j = 1, \cdots, n), \tag{6ν}$$

$(i = 1, \cdots, n; \nu = 0, 1, 2, \cdots)$. Here $d_{ik} \equiv 0$ for $k < 0$ by definition, and the $f_{i\nu}$ are polynomials in $d_{j1}, d_{j2}, \cdots, d_{j\nu}$ $(j = 1, \cdots, n)$ with coefficients regular in d_{j0} $(j = 1, \cdots, n)$.

In the classical case, $s_i = 0$ $(i = 1, \cdots, n)$ and the conditions (6ν) are recursive; the coefficients d_{ij} for $j = 2, 3, \cdots$ are known explicitly as regular functions of the n parameters $d_{10}, \cdots, d_{n0}$. By Cauchy's majorant method it is then proved ([6]) that for every $r < r_1$, an r_3 can be taken so small that (5) converge for $|z| \leqq r$ and satisfy (3b) whenever $|d_{i0}| < r_3$ $(i = 1, \cdots, n)$ holds.

In the general case, however, the conditions (6ν) are not recursive. Note that by (2) we have in general excluded the Fuchsian theory of regular singular points $(s_i = 1, g_i \equiv 0; i = 1, \cdots, n)$, but have not excluded those regular singular points for which some f_i contains a factor z. (Nor have we excluded those irregular singular points which satisfy the necessary conditions ([1], p. 417) for the existence of $n - s$ solutions "regular" in the sense of [1].)

Suppose now that some s_i is positive, so that the system (6ν) is non-recursive.

When the f_i are linear in the w_j (i.e., $g_i \equiv 0$), the conditions (6ν) are linear. Consequently the theory of infinite matrices may be applied. By

using the resolvent series (C. Neumann), Perron ([5]) and Lettenmeyer ([3]) showed that the system (6ν) $(\nu = 0, 1, \cdots)$ admits of at least $n - s$ linearly independent solutions, each of which is such that the corresponding solution (5) of (3) converges for all $|z| \leqq r_1$.

In the general case, the conditions (6ν) are neither recursive nor linear. Our method of proof will be to effect a change of variables such that Wintner's fixed-point theorem for analytic (but not in general completely continuous) mappings in a separable Hilbert space is applicable. This theorem allows us to conclude that for a sufficiently large integer N the components of the vector representing the fixed point will be regular in the d_{ij} $(i = 1, \cdots, n;$ $j = 1, \cdots, N)$ and any parameters in which the f_i are given regular. (The algebroid dependence occurs only among the d_{ij} $(i = 1, \cdots, n; j = 1, \cdots, N)$.) Since the (modified) coefficients of the power series (5) obtained in this way are components of a vector in the Hilbert space we can easily prove the convergence of the solution (5).

Our change of variables is

$$\xi_{ik} = k d_{ik} r^k, \tag{7}$$

$(i = 1, \cdots, n; k = 1, 2, \cdots)$, $\xi_{i0} = d_{i0}$, which differs slightly from that of Lettenmeyer: our proof of the boundedness of the matrix of the linear terms of the $f_{i\nu}$ in (6ν) roughly parallels his proof, but by using (7) we can conclude directly (without explicit estimates) that the non-linear terms do form a mapping in Hilbert space. In fact, explicit estimates of his type would imply that the power series (5) converge only in the open circle $|z| < r$; but (7) and Schwarz's inequality imply that the series (5) converge absolutely on the closed circle

$$|z| \leqq r. \tag{8}$$

Now (8) allows the Parseval relation for the g_i to be applied at $|z| = r$, which produces the desired result.

We shall also generalize Theorem 1 as follows:

THEOREM 2. *Let* $f_i(z, w_1, \cdots, w_n)$, $(i = 1, \cdots, n)$, *be as in Theorem* 1. *Let* $b_{ik}(z)$, $(i, k = 1, \cdots, n)$, *be regular on some neighborhood of* $z = 0$ *and convergent on* (3a); *furthermore, assume that* $\det \| b_{ik}(z) \| \not\equiv 0$, *and that* $\det \| b_{ik}(z) \|$ *has an s-fold zero at* $z = 0$ *with* $s < n$. *Then the system of differential equations*

$$\sum_{k=1}^{n} b_{ik}(z) w_k'(z) = f_i(z, w_1, \cdots, w_n), \tag{9}$$

$(i=1,\cdots,n)$, has at least an $(n-s)$-parameter algebroid family of solutions regular at $z=0$. The solutions will have algebroid dependence on any additional parameters in which the f_i are given regular.

2. Definition. We shall say that (3) has an *m-parameter algebroid family* of solutions regular at the origin when the following conditions are satisfied:

(i) For every positive $r<r_1$ there exist a positive $r_3<r_2$, $2m$ non-negative integers i_k, j_k, $(k=1,\cdots,m)$, a positive integer N and n^2N^2 (complex) constants p_{ijkl}, $(i,k=1,\cdots,n;j,l=1,\cdots,N)$, such that for any choice of complex numbers μ_k, $(k=1,\cdots,m)$, satisfying

$$|\mu_k|<r_3, \qquad (k=1,\cdots,m), \tag{10}$$

there exists a power series solution (5) of (3), absolutely convergent on (8), satisfying (3b), and such that if we put

$$\mu_{ij}=\sum_{k=1}^{n}\sum_{l=1}^{N}p_{ijkl}d_{kl}, \qquad (i=1,\cdots,n;j=1,\cdots,N). \tag{10'}$$

then

(ii) the transformation (10′) is non-singular, and

(iii) *all the initial conditions $d_{ij}\equiv w_i^{(j)}(0)$, $(i=1,\cdots,n;j=1,2,\cdots)$ are algebroid functions of the m parameters μ_k, where*

$$\mu_k=\mu_{i_kj_k}, \qquad (k=1,\cdots,m). \tag{11}$$

If the f_i contain q additional parameters, then the μ_{ij} of (10′) depend linearly on them (as well as the d_{ik}), which by (11) implies the algebroid dependence of the solution on these $m+q$ parameters.

3. Proof of Theorem 1. Substitute (5) into $g_i(z,w_1,\cdots,w_n)$ and expand formally,

$$g_i=\sum_{\nu=0}^{\infty}g_{i\nu}(d_{kl})z^\nu \qquad (i=1,\cdots,n), \tag{12}$$

to define the polynomials $g_{i\nu}=g_{i\nu}(d_{k1},\cdots,d_{k\nu};k=1,\cdots,n)$ whose coefficients are uniformly convergent power series in the d_{k0} $(k=1,\cdots,n)$. Also, let

$$a_{ij}(z)=\sum_{k=0}^{\infty}a_{ijk}z^k, \qquad (i,j=1,\cdots,n). \tag{13}$$

Let M_1 be the largest maximum modulus of the g_i on (3a), (3b), and M_2 the largest for the a_{ij} on (3a). Let r be fixed such that $0 < r < r_1$. By the remark following (8), whenever (5) converges on (8) and satisfies (3b) there,

$$\sum_{k=0}^{\infty} |\, g_{ik} \,|^2 r^{2k} \leqq M_1^2, \qquad (i = 1, \cdots, n). \tag{14}$$

Since the g_i are, by (4), power series containing no terms of degree lower than the second, replacing the w_k in (3) by $\mu^1 w_k$, $(k = 1, \cdots, n)$, replaces g_i in (12) by $(\mu^2)^1 g_i(d_{kl}, \mu)$, where ${}^1g_i(d_{kl}, \mu)$ is a power series in μ. If this change of scale is made, (12) and (13) inserted in (16ν), and the result divided by μ, we have the equivalent conditions

$$(\nu - s_i + 1) d_{i\,\nu-s_i+1} = \sum_{j=1}^{n} \sum_{\lambda=0}^{\nu} a_{ij\,\nu-\lambda} d_{j\lambda} + \mu^1 g_i(d_{kl}, \mu). \tag{15$i\nu$}$$

Split off the $nN + s$ equations (15$i\nu$) for $\nu = 0, 1, \cdots, N + s_i - 1$, $(i = 1, \cdots, n)$, and call them (15a); call the remaining infinite system of equations (15b).

Now when (7) is used, (15$i\nu$) become

$$\xi_{i\,\nu-s_i+1} = F_{i\nu}(\xi_{kl}; l = 0, \cdots, \nu) + \mu G_{i\nu}, \tag{15$i\nu'$}$$

where (letting $1/\lambda = 1$ at $\lambda = 0$)

$$F_{i\nu} = r/r^{s_i} \sum_{j=1}^{n} \sum_{\lambda=0}^{\nu} (a_{ij\,\nu-\lambda} r^{\nu-\lambda}/\lambda) \xi_{j\lambda}; \tag{16a}$$

$$G_{i\nu} = (r/r^{s_i})[{}^1g_{i\nu}(d_{kl}, \mu) r^{\nu}] = (r/r^{s_i})[{}^1g_{i\nu}(\xi_{kl}/k! r^k, \mu) r^{\nu}]. \tag{16b}$$

To consider (15b′) put $\nu - s_i = \alpha$ in (15$i\nu'$), (16a), and (16b), and require $\alpha \geqq N$ (N is to be fixed later). We then have

$$\xi_{i\,\alpha+1} = F_{i\,\alpha+s_i} + \mu G_{i\,\alpha+s_i}(\mu), \qquad (\alpha \geqq N). \tag{15b′}$$

Let r_3 be such that $0 < r_3 < 2r_3 < \min(r_1, r_2)$ and require that

$$|\, \xi_{i0} \,| < 2r_3 < \min(r_1, r_2). \tag{17}$$

Suppose also (subject to later verification) that (15a′) can be solved in such a manner that

$$|\, \xi_{ik} \,| < 2\epsilon, \qquad (k = 1, \cdots, N), \tag{18}$$

where $0 < \epsilon \leqq r_3$. Since the right hand sides of (15b′) contain no ξ_{ij} with j greater than $\alpha + s_i$, the functions $F_{i\,\alpha+s_i}$ and $G_{i\,\alpha+s_i}$ are "analytic functions

of the infinitely many variables $\xi_{i\,N+1}, \xi_{i\,N+2}, \cdots$ $(i=1,\cdots,n)$" in the sense of [7]. The remaining ξ_{ik} $(k=0,1,\cdots,N)$ are for the present to be regarded as parameters. Note that by (16a), (16b), and the remark following (12), the F_{ik} and G_{ik} converge uniformly in these parameters when the parameters are restricted by (17) and (18).

Let $\boldsymbol{M}_{2r_3}$ and $\boldsymbol{N}_{2\epsilon}$ be the regions (17) and (18), and define $\boldsymbol{M}_{r_3}$ and $\boldsymbol{N}_{\epsilon}$ similarly. Let $\boldsymbol{k}_\Delta$ represent the open disk $|\mu|<\Delta$ and $\boldsymbol{K}_\delta$ the sphere

$$\sum_{i=1}^{n}\sum_{k=N+1}^{\infty}|\xi_{ik}|^2<\delta^2. \tag{19}$$

Fix δ such that

$$\delta<3^{\frac{1}{2}}(r_2-2r_3)/\pi; \tag{20}$$

note that decreasing r_3 (which must be done later) does not disturb the inequality (20). At first suppose merely that $\epsilon\leqq r_3<\delta$, so that whenever (17), (18), and (19) hold, the inequality

$$|\xi_{ik}|<2\delta \tag{21}$$

holds for all i, k.

If we can show that for some fixed Δ, say $\Delta_1>0$, there exist numbers ω_N and ω such that, on the product set $\boldsymbol{K}_\delta\times\boldsymbol{M}_{2r_3}\times\boldsymbol{N}_{2\epsilon}$,

$$\sum_{i=1}^{n}\sum_{k=N}^{\infty}|F_{i\,k+s_i}|^2<\omega_N, \tag{22}$$

$$\sum_{i=1}^{n}\sum_{k=N}^{\infty}|G_{i\,k+s_i}|^2<\omega, \tag{23}$$

and that (keeping δ fixed) $\omega_N\to 0$ as $N\to\infty$, then by Theorem IIa of [7] there exist an N and a $\Delta<\Delta_1$ such that (15b) has exactly one solution $\xi_{kj}(\mu)$, analytic on $\boldsymbol{k}_\Delta$ and satisfying (19), for each point of $\boldsymbol{M}_{2r_3}\times\boldsymbol{N}_{2\epsilon}$. (This is true because by letting Δ be sufficiently small, not only ω_N but also $|\mu|^2\omega$ can be made sufficiently small in relation to the fixed δ.) This solution ξ_{kj} $(k=1,\cdots,n;\ j=N+1,N+2,\cdots)$ is analytic in all parameters on $\boldsymbol{k}_\Delta\times\boldsymbol{M}_{r_3}\times\boldsymbol{N}_{\epsilon}$; it is also analytic (for a halved radius of convergence) in any other parameters in which the f_i (and so the F's and G's) are given analytic.

We first consider the validity of (22). By (16a) and (21),

$$|F_{i\,a+s_i}|<2r\delta/r^{s_i}\sum_{j=1}^{n}\sum_{\lambda=0}^{a+s_i}|a_{ij\,a+s_i-\lambda}|\,r^{a+s_i-\lambda}/\lambda. \tag{24}$$

Now by the choice of r, $\theta=r/r_1<1$. We next use Cauchy's coefficient estimate for (13) to establish that

$$|F_{i\,a+s_i}|<2nr\delta M_2/r^{s_i}\sum_{\lambda=0}^{a+s_i}\theta^{a+s_i-\lambda}/\lambda. \tag{25}$$

By splitting the sum in (25), using Schwarz's inequality on the first half and majorization by a geometric series on the second, we find

$$\sum_{\lambda=0}^{\alpha+s_i} \theta^{\alpha+s_i-\lambda}/\lambda < c_1\theta^{(\alpha+s_i)/2} + 2c_2/(\alpha+s_1-1), \tag{26}$$

where $c_1 = [(1+\pi^2/6)/(1-\theta)]^{\frac{1}{2}}$ and $c_2 = 1/(1-\theta)$. By (25), (26), and Minkowski's inequality, we can take for ω_N in (22)

$$\omega_N = n(2nr\delta M_2/r^{s_1})^2\Omega_N, \tag{27}$$

where $\Omega_N = [(c_1\theta^{N/2}/(1-\theta^{\frac{1}{2}})^{\frac{1}{2}}) + 2c_2(1/(N-1)^2 + \cdots)^{\frac{1}{2}}]^2$. Consequently, we can choose N so large that ω_N is as small as required.

We have assumed $\epsilon \leqq r_3 < \delta$. Now that N is fixed, we may further diminish r_3 if necessary, in order to ensure that

$$4nN\epsilon^2 \leqq 4nNr_3^2 < \delta. \tag{28}$$

We shall next prove (23), under the hypotheses (17), (18), (19). By (28), the hypotheses (18), (19) imply

$$\sum_{j=1}^{n}\sum_{k=1}^{\infty} |\xi_{jk}|^2 < 2\delta^2. \tag{29}$$

This holds a fortiori if the first summation sign is removed. Hence by Schwarz's inequality and the fact that $\sum 1/k^2 = \pi^2/6$,

$$\sum_{k=1}^{\infty} |d_{jk}|\, r^k = \sum_{k=1}^{\infty} |\xi_{jk}|/k \leqq \pi\delta/3^{\frac{1}{2}}, \tag{30}$$

which proves that the series (5) converge absolutely on (8). Furthermore, by (17), (20), and (30),

$$|w_i(z)| \leqq |d_{i0}| + |\sum_{k=1}^{\infty} d_{ik}z^k| < r_2. \tag{31}$$

In proving (23), we need only (by the maximum modulus theorem for the partial sums of (23), as stated in Littlewood [2], p. 238) consider the functions $g_{i\nu}(\mu)$ of (12), (14), and (16b) at $|\mu| = 1$, as will now be shown. Equation (31) (which continues to hold when the d_{ik} are each multiplied by $e^{i\phi}, \phi > 0$) shows that for any choice of the coefficients of (5) satisfying (17), (18), and, above all, (19), the power series (5) can be inserted in the g_i and expanded as in (12). Then (14) holds, and continues to hold, if the first $N + s_i - 1$ terms are omitted. Comparing this with (16b), we see that on $|\mu| = 1$ we may put

$$\omega = r^2M_1^2\sum_{j=1}^{n} r^{-2s_j}. \tag{32}$$

But then by the argument of [2], (23) holds for all $|\mu| \leqq 1$.

Thus we have not only shown that the system of conditions (15b′) possesses solutions $\xi_{jk} = \xi_{jk}(\mu, \xi_{j0}, \cdots, \xi_{jN}; r)$, $(k = N+1, N+2, \cdots)$, analytic on $\boldsymbol{k}_\Delta \times \boldsymbol{M}_{r_3} \times \boldsymbol{N}_\epsilon$, but in (31) we have shown that the solutions (5) resulting from (15b′) converge on (8) and satisfy (3b), provided only that (15a′) and (15b′) can be satisfied simultaneously for values of ξ_{jk}, $(k = 0, 1, \cdots, N)$, lying in $\boldsymbol{M}_{r_3} \times \boldsymbol{N}_\epsilon$.

First, fix μ on $\boldsymbol{k}_\Delta$. Then the solutions of (15b′) just established can be rewritten by means of (7) in the original variables:

$$d_{j\,N+1} = \eta_{j1}(d_{j0}, \cdots, d_{jN}), \cdots, d_{j\,N+\sigma-1} = \eta_{j\,\sigma-1}(d_{j0}, \cdots, d_{jN}) \tag{33}$$

$$d_{j\,N+\rho} = \eta_{j\rho}(d_{j0}, \cdots, d_{jN}), \qquad (\rho = \sigma, \sigma+1, \cdots) \tag{34}$$

where $\sigma = \max s_i$, $(i = 1, \cdots, n)$. The functions η_{jk} are analytic in their $N+1$ variables. Since the solution (33)-(34) is unique and (3) admits the solution $w_i \equiv 0$, $(i = 1, \cdots, n)$, we know that $\eta_{jk}(0, 0, \cdots, 0) = 0$ for all j, k.

The $\sum_{i=1}^{n}(N + s_i) = nN + s$ equations (15a′) have the form

$$\begin{gathered} 0 = f_{j0}(d_{j0}), \cdots, 0 = f_{j\,s_j-1}(d_{j0}, \cdots, d_{j\,s_j-1}) \\ d_{j1} = f_{js_j}(d_{j0}, \cdots, d_{js_j}), \cdots, d_{jN} = f_{j\,N+s_j-1}(d_{j0}, \cdots, d_{j\,N+s_j-1}). \end{gathered} \tag{35}$$

According to the definition of σ, none of the equations (35) contain variables which do not occur in the $n(\sigma-1)$ equations (33). Furthermore, $f_{jk}(0, \cdots, 0) = 0$, $(k = 0, \cdots, N + s_j - 1)$. Considering (33) and (35) as a single system, we then have $nN + s + n\sigma - n$ analytic implicit equations, vanishing at the origin, for the $n(N+\sigma)$ unknowns $d_{j0}, \cdots, d_{j\,N+\sigma-1}$, $(j = 1, \cdots, n)$. Consequently, by a theorem of Weierstrass (as stated in [4], p. 132), at least $n(N+\sigma) - (nN + s + n\sigma - n) = n - s$ linear combinations of the unknowns can be left as parameters, say the μ_k of (11), with $m \geqq n - s$; and the remaining unknowns can be found as algebroid functions of them. Moreover, the transformation (10′) is non-singular, and since these algebroid functions disappear at $\mu_k = 0$, $(k = 1, \cdots, m)$, the equations (17) and (18) can be satisfied (with r_3 and ϵ replaced by $\frac{1}{2}r_3$ and $\frac{1}{2}\epsilon$) by decreasing the value of r_3 which appears in (10).

4. Proof of Theorem 2. Let $c_{ik}(z)$, $d_{ik}(z)$, $(i, k = 1, \cdots, n)$, be power series regular near $z = 0$ and such that $\det \| c_{ik} \|$, $\det \| d_{ik} \|$ have no zeros in some neighborhood of the origin. In (9) put $w_k = \sum_{j=1}^{n} c_{ik}(z) v_k$; then (9) becomes a system (9′) of the same type for v_k. If f_i is given by (4), then

b_{ik} is replaced by $\sum_{l=1}^{n} b_{il}c_{lk}$ and a_{ik} replaced by $\sum_{j=1}^{n}(a_{ij}b_{jk} - b_{ij}c_{jk}')$. The new g_i are of the second degree or higher in the v_k. Now multiply the k-th equation of (9′) by $d_{ik}(z)$ and sum over k $(k = 1, \cdots, n)$; the result is a system (9″) of the form (9) again, but with b_{ik} replaced by $e_{ik} = \sum_{j=1}^{n}\sum_{l=1}^{n} d_{ij}(z)b_{jl}(z)c_{lk}(z)$. Lettenmeyer [3] has shown that $d_{ij}(z)$ and $c_{lk}(z)$ of the type mentioned can be so chosen that $e_{ik} = z^{s_i}\delta_{ik}$ $(\delta_{ik} = 0, i \neq k; = 1, i = k)$, where s_i are integers satisfying (2). Theorem 2 then follows as an immediate corollary of Theorem 1.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] E. L. Ince, *Ordinary differential equations*, 1927, p. 417.

[2] J. E. Littlewood, *Lectures on the theory of functions*, 1944, p. 238.

[3] F. Lettenmeyer, "Über die an einer Unbestimmtheitsstelle regulären Lösungen eines Systems homogener linearer Differentialgleichungen," *Sitzungsberichte München* (1926), pp. 287-307.

[4] W. F. Osgood, *Lehrbuch der Funktionentheorie*, 2nd ed. (1929), vol. II_1, p. 132.

[5] O. Perron, "Über diejenigen Integrale linearer Differentialgleichungen welche sich an einer Unbestimmtheitsstelle bestimmt verhalten," *Mathematische Annalen*, vol. 70 (1911), pp. 1-32.

[6] H. Poincaré, *Les méthodes nouvelles de la mécanique céleste*, vol. I (1892), p. 55.

[7] A. Wintner, "Zur Analyse im Hilbertschen Raume," *Mathematische Zeitschrift*, vol. 28 (1928), pp. 457-470.

REPRESENTATIONS OF SEMISIMPLE LIE GROUPS IV.*

By HARISH-CHANDRA.

---

1. Introduction. In an earlier paper of this series, I have given a method of constructing all irreducible quasi-simple representations of a connected semisimple group G (at least in case G has a finite-dimensional faithful representation) up to infinitesimal equivalence (see Theorem 4 of [5(d)]. However this method has one serious drawback. It can happen very well that the representation so obtained is infinitesimally unitary without being equivalent to a unitary representation. In order to make it unitary one has to introduce a new scalar product in a dense subspace of the representation space in some more or less artificial fashion and then complete this subspace with respect to the new norm. It would of course be much better if the unitary representations could be obtained in a more natural manner especially since it is often a rather difficult matter to decide whether a given representation is infinitesimally unitary or not. Suppose, for a moment, that G is a linear group and A is a Cartan subgroup of G. Then A can be written as a direct product of two subgroups A_- and A_+ of which the former is compact while the latter is isomorphic to a vector space. We say that A is incompact if A_+ has the maximum possible dimension. Now in the above-mentioned construction it was quite essential that one should start from an incompact A. However the study of simple examples (such as that of the 2×2 real unimodular group [1, 5(e)]) is enough to convince one of the necessity of searching for alternative 'natural' methods of construction which start out from other types of Cartan subgroups. The work of Gelfand and Graev [4] on the Plancherel formula of the $n \times n$ real unimodular group points to the same conclusion. In a recent note [5(f)] I have given such a 'natural' mode of construction based on the concept of induced representations (see Mackey [6]), which is applicable to all Cartan subgroups. By a simple argument the main problem can be reduced to the case when A is compact and therefore it is this particular case which requires intensive study. The object of the present paper is to deal with the algebraic aspects of this case, the function-theoretic aspects being left to a subsequent paper. Assuming that G is simple and its first Betti number is 1, we shall

---

* Received July 8, 1955.

give a quite general method of constructing a series of irreducible unitary representations of G parameterised by the characters of A. However if the first Betti number is zero (and G is not compact) this method does not give anything at all [5(g)]. (It seems likely that this is not due to a fault of the method but rather reflects an essential difference in the two cases with regard to representation theory. It is noteworthy that a similar difference has been observed by E. Cartan [3(d)] in his theory of bounded homogeneous symmetric domains. There also exactly those simple groups come into play whose first Betti number is one.) The representations obtained in this manner bear a very close resemblance, in their algebraic structure, to finite-dimensional representations. As we shall see in another paper, they are also somewhat remarkable in so far as their matrix coefficients are square-integrable on the group [5(f), (g)]. In the last section we obtain a formula (see Corollary to Lemma 21) for the multiplicities of the various weights of these representations.

Some of the results of this paper have been announced in a short note [5(g)].

2. Preliminary lemmas. Let R be the field of real numbers and V a vector space over R of finite dimension. We shall say that V is ordered if there is given an order in V satisfying the following condition. If u, v are two elements in V and a is a positive real number then $u+v>0$ and $au>0$ whenever $u>0$, $v>0$. An element $u \in V$ will be called positive (with respect to the given order) if $u>0$. Let $\mathfrak{F}$ be the space of all linear functions on V. Then there are two simple ways of introducing an order in $\mathfrak{F}$ which depend on the choice of a base either in V or in $\mathfrak{F}$. Let $H_1, \cdots, H_l$ be a base for V. We define a function $\lambda \in \mathfrak{F}$ to be positive if $\lambda \neq 0$ and $\lambda(H_i)>0$, i being the least index such that $\lambda(H_i) \neq 0$. This is called the lexicographic order in $\mathfrak{F}$ corresponding to the base $(H_1, \cdots, H_l)$. Similarly let $(\lambda_1, \cdots, \lambda_l)$ be a base for $\mathfrak{F}$. Then every $\lambda \in \mathfrak{F}$ can be written as $\lambda = c_1\lambda_1 + \cdots + c_l\lambda_l$ $(c_j \in R)$. We now define λ to be positive if $\lambda \neq 0$ and $c_i>0$ where i is the least index such that $c_i \neq 0$. The order so defined is called the lexicographic order in $\mathfrak{F}$ corresponding to the base $(\lambda_1, \cdots, \lambda_l)$.

Let $\mathfrak{g}$ be a semisimple Lie algebra over the field C of complex numbers and let $\mathfrak{h}$ be a Cartan[1] subalgebra of $\mathfrak{g}$. Let $\mathfrak{h}_R$ denote the subset of those elements of $\mathfrak{h}$ at which every root α (of $\mathfrak{g}$ with respect to $\mathfrak{h}$) takes a real value. Then $\mathfrak{h}_R$ is a vector space over R and if $H_1, \cdots, H_l$ is a base for $\mathfrak{h}_R$

[1] We use the standard terminology of the theory of semisimple Lie algebras (see for example [5(b)]).

over R, it is also a base for $\mathfrak{h}$ over C. We say that a linear function λ on $\mathfrak{h}$ is real if $\lambda(H)$ is real for $H \in \mathfrak{h}_R$. Then the set $\mathfrak{F}_R$ of all such real functions may be identified in the obvious way with the space of all linear functions on $\mathfrak{h}_R$. Hence corresponding to the base $(H_1, \cdots, H_l)$ we get a lexicographic order in $\mathfrak{F}_R$. Similarly if $\alpha_1, \cdots, \alpha_l$ are l linearly independent roots they form a base for $\mathfrak{F}_R$ and we obtain the corresponding lexicographic order in $\mathfrak{F}_R$.

Let π be a representation of $\mathfrak{g}$ on a vector space V over C. A linear function Λ on $\mathfrak{h}$ is called a weight of π if there exists an element $\psi \neq 0$ in V such that $\pi(H)\psi = \Lambda(H)\psi$ for all $H \in \mathfrak{h}$. A weight Λ is called extreme if it is impossible to find a root α such that both $\Lambda + \alpha$ and $\Lambda - \alpha$ are also weights of π. It is well known (see Weyl [7]) that if $\dim V$ is finite every weight Λ is real.

For each root α choose an element $X_\alpha \neq 0$ in $\mathfrak{g}$ such that $[H, X_\alpha] = \alpha(H)X_\alpha$ for all $H \in \mathfrak{h}$. Then $H_\alpha = [X_\alpha, X_{-\alpha}]$ lies in $\mathfrak{h}$ and we may suppose that X_α is so chosen that $\alpha(H_\alpha) = 2$. Let $\mathfrak{B}$ be the universal enveloping algebra of $\mathfrak{g}$. We identify representations of $\mathfrak{g}$ with those of $\mathfrak{B}$ in the usual way.

LEMMA 1. *Let π be a representation of $\mathfrak{g}$ on V and let $\psi \neq 0$ be an element belonging*[1] *to the weight Λ. Suppose α is a root such that $\pi(X_\alpha^r)\psi = \pi(X_{-\alpha}^r)\psi = 0$ for some integer $r \geqq 1$. Then $\Lambda(H_\alpha)$ is an integer and $\Lambda - k\alpha$ is a weight of π for all integers k such that*

$$\min(0, \Lambda(H_\alpha)) \leqq k \leqq \max(0, \Lambda(H_\alpha)).$$

In particular $\Lambda - \Lambda(H_\alpha)\alpha$ is a weight of π. Moreover if $\pi(X_\alpha)\psi = 0$, $\Lambda(H_\alpha)$ is equal to the smallest integer $r \geqq 0$ such that $\pi(X_{-\alpha}^{r+1})\psi = 0$.

The proof of the first two statements is substantially the same as that of Lemma 11 of [5(b)]. Now suppose $\pi(X_\alpha)\psi = 0$. Let s be the least non-negative integer such that $\pi(X_{-\alpha}^{s+1})\psi = 0$. Then if $\psi_{-1} = 0$, $\psi_0 = \psi$ and $\psi_k = \pi(X_{-\alpha}^k)\psi$ $(k \geqq 1)$ the argument of Lemma 11 of [5(b)] shows that

$$\pi(X_\alpha)\psi_k = k(\Lambda(H_\alpha) - k + 1)\psi_{k-1} \qquad (k \geqq 0).$$

Putting $k = s + 1$ we get $\Lambda(H_\alpha) = s$ and this proves the last part of assertion.

Let λ be a linear function on $\mathfrak{h}$. We say that an element $b \in \mathfrak{B}$ is of rank λ if $[H, b] = \lambda(H)b$ for all $H \in \mathfrak{h}$. Elements belonging to different ranks are linearly independent and if $\mathfrak{B}_\lambda$ is the set of all elements in $\mathfrak{B}$ of rank λ, $\mathfrak{B}$ is the direct sum of $\mathfrak{B}_\lambda$ for all λ [5(b), Lemma 2]. By the component of z of rank λ $(z \in \mathfrak{B})$ we mean the component of z in $\mathfrak{B}_\lambda$ in this direct sum decomposition.

Now suppose $\mathfrak{F}_R$ has been ordered in some way.

LEMMA 2. *Let π be a representation of $\mathfrak{B}$ on a vector space V. Suppose Λ is a linear function on $\mathfrak{h}$ and $\psi \neq 0$ a vector in V such that:*

(1) $\pi(X_\alpha)\psi = 0$ *for every positive root* α;

(2) $\pi(H)\psi = \Lambda(H)\psi$ *for all* $H \in \mathfrak{h}$;

(3) $V = \pi(\mathfrak{B})\psi$.

Then if $\alpha_1, \cdots, \alpha_r$ are all the distinct positive roots of $\mathfrak{g}$ every weight of π is of the form $\Lambda - (m_1\alpha_1 + \cdots + m_r\alpha_r)$ where $m_1, \cdots, m_r$ are nonnegative integers. Moreover if for each α_i we can find an integer $n_i \geqq 0$ such that $\pi(X_{-\alpha_i}^{n_i+1})\psi = 0$ then dim V *is finite and π is irreducible.*

Since $\mathfrak{g}$ is spanned by $(X_{\alpha_1}, \cdots, X_{\alpha_r}, \mathfrak{h}, X_{-\alpha_1}, \cdots, X_{-\alpha_r})$ it follows from Lemma 1 of [5(a)] and our assumptions above that V is spanned by elements of the form $\pi(X_{-\alpha_1}^{m_1} X_{-\alpha_2}^{m_2} \cdots X_{-\alpha_r}^{m_r})\psi$. In view of Lemma 10 of [5(b)] our first assertion now follows immediately. Moreover it is clear that $C\psi$ is the space of all elements in V which belong to the weight Λ.

Now suppose $\pi(X_{-\alpha_i}^{n_i+1})\psi = 0$, $1 \leqq i \leqq r$. Then we claim that V is spanned by $\pi(X_{-\alpha_1}^{\nu_1} \cdots X_{-\alpha_r}^{\nu_r})\psi$, $(0 \leqq \nu_i \leqq n_i, 1 \leqq i \leqq r)$. For let U be the space spanned by these elements. It would be sufficient to show that $\pi(X_{-\alpha_1}^{m_1} \cdots X_{-\alpha_r}^{m_r})\psi \in U$ for all $m_1, \cdots, m_r \geqq 0$. We shall do this by induction on $M = m_1 + \cdots + m_r$. If $M = 0$ the statement is obviously true and so we may suppose $M > 0$. Also we may assume that $m_i > n_i$ for some i. Let j be the largest index such that $m_j > n_j$. If $j = r$, $\pi(X_{-\alpha_r}^{m_r})\psi = 0$ and so our statement is again true. Hence suppose $j < r$. Put

$$z = X_{-\alpha_1}^{m_1} \cdots X_{-\alpha_{j-1}}^{m_{j-1}}, \qquad w = X_{-\alpha_{j+1}}^{m_{j+1}} \cdots X_{-\alpha_r}^{m_r}$$

($z = 1$ if $j = 1$). Then $zX_{-\alpha_j}^{m_j} w = z[X_{-\alpha_j}^{m_j}, w] + zwX_{-\alpha_j}^{m_j}$ and since $m_j > n_j$, $\pi(X_{-\alpha_j}^{m_j})\psi = 0$. Therefore

$$\phi = \pi(X_{-\alpha_1}^{m_1} \cdots X_{-\alpha_r}^{m_r})\psi = \pi(z[X_{-\alpha_j}^{m_j}, w])\psi.$$

But $X_{-\alpha_1}, \cdots, X_{-\alpha_r}$ span a subalgebra of $\mathfrak{g}$ and therefore from Lemma 1 of [5(a)] $z[X_{-\alpha_j}^{m_j}, w]$ can be written as a linear combination of $X_{-\alpha_1}^{m_1'} \cdots X_{-\alpha_r}^{m_r'}$ with $m_1' + \cdots + m_r' < M$. Hence by our induction hypothesis $\phi \in U$. This proves that $U = V$ and therefore dim V is finite and V is fully reducible under π. Hence there exists a subspace of V' which is invariant and irreducible under $\pi(\mathfrak{B})$ and which contains a nonzero vector ψ' belonging to the weight Λ. But then, as we have seen above $\psi' = c\psi$ $(c \in C)$ and therefore $V' = \pi(\mathfrak{B})\psi' = V$. This proves that π is irreducible.

We shall need the following lemma due to E. Cartan [3(a)].

LEMMA 3 (Cartan). *Suppose π is an irreducible finite-dimensional representation of $\mathfrak{g}$ and Λ_0 is its highest weight. Then if W is the Weyl*[1] *group of $\mathfrak{g}$ (with respect to $\mathfrak{h}$) $s\Lambda_0$ ($s \in W$) are exactly all the extreme weights of π.*

We sketch the proof for the sake of completeness. It follows from the argument of Lemma 13 of [5(b)] that if Λ is an extreme weight of π so is also $s\Lambda$ for every $s \in W$. Since the highest weight is obviously an extreme weight, $s\Lambda_0$ ($s \in W$) are all extreme weights. Conversely suppose Λ is an extreme weight of π. Let Λ_1 be the highest one among $s\Lambda$ ($s \in W$). Then Λ_1 is also an extreme weight and it would be enough to prove that $\Lambda_1 = \Lambda_0$. Let α be a positive root and s_α the corresponding Weyl reflexion. Then $s_\alpha \Lambda_1 = \Lambda_1 - \Lambda_1(H_\alpha)\alpha \leqq \Lambda_1$. Hence $\Lambda_1(H_\alpha) \geqq 0$. We claim that $\Lambda_1 + \alpha$ is not a weight of π. For if it were we could conclude from Lemma 1 (since $\pi(X_\alpha)$ and $\pi(X_{-\alpha})$ are both nilpotent) that $\Lambda_1 + \alpha - \nu\alpha$ is a weight for $0 \leqq \nu \leqq \Lambda_1(H_\alpha) + 2$. This would imply however that $\Lambda_1 + \alpha$ and $\Lambda_1 - \alpha$ are both weights, contradicting the fact that Λ_1 is extreme. Let $\psi \neq 0$ be an element in V belonging to the weight Λ_1. Since $\Lambda_1 + \alpha$ is not a weight, $\pi(X_\alpha)\psi = 0$ and it follows from Lemma 2 that Λ_1 is the highest weight of π. Therefore $\Lambda_1 = \Lambda_0$ and Lemma 3 is proved.

A set $(\alpha_1, \cdots, \alpha_l)$ of roots of $\mathfrak{g}$ is called a fundamental system if it is a linearly independent set and if every root α can be written in the form $\alpha = d_1\alpha_1 + \cdots + d_l\alpha_l$ where d_i are integers which are either all nonnegative or all nonpositive.

LEMMA[2] 4. *Let $\mathfrak{m}$ be a subalgebra of $\mathfrak{g}$ such that $\mathfrak{h} \subset \mathfrak{m}$ and for each root α either X_α or $X_{-\alpha}$ is in $\mathfrak{m}$. Then it is possible to choose a fundamental system $(\alpha_1, \cdots, \alpha_l)$ of roots such that $X_{\alpha_i} \in \mathfrak{m}$, $1 \leqq i \leqq l$.*

Let Γ be the radical of $\mathfrak{m}$ and let $X \to X^*$ denote the natural mapping of $\mathfrak{m}$ on $\mathfrak{m}^* = \mathfrak{m}/\Gamma$. Then $\mathfrak{m}^*$ is a semisimple Lie algebra and $\mathfrak{h} = (\mathfrak{h} + \Gamma)/\Gamma$ is an abelian subalgebra of $\mathfrak{m}^*$. Let $X \to adX$ denote the adjoint representation of $\mathfrak{g}$. Then if $H \in \mathfrak{h}$, adH is a semisimple[3] linear transformation of $\mathfrak{g}$. Hence if ad^*H^* denotes the linear mapping of $\mathfrak{m}^*$ corresponding to H^* under the adjoint representation of $\mathfrak{m}^*$, it is clear that ad^*H^* is also semisimple. Therefore $\mathfrak{h}^*$ can be extended to a Cartan subalgebra $\mathfrak{h}_1^*$ of $\mathfrak{m}^*$. We introduce some lexicographic order among the roots of $\mathfrak{m}^*$ with respect to $\mathfrak{h}_1^*$. For every root γ^* of $\mathfrak{m}^*$ choose an element $X_{\gamma^*}^* \neq 0$ in $\mathfrak{m}^*$ such that $[H^*, X_{\gamma^*}^*] = \gamma^*(H^*)X_{\gamma^*}^*$ for all $H^* \in \mathfrak{h}_1^*$. Let $\Gamma_1^* = \mathfrak{h}_1^* + \sum_{\gamma^* > 0} CX_{\gamma^*}$ and

[2] A very similar result has been obtained independently by A. Borel.

[3] This means that $\mathfrak{g}$ is fully reducible under adH.

let Γ_1 be the complete inverse image of Γ_1^* in $\mathfrak{m}$. It is clear that Γ_1^* is a solvable subalgebra of $\mathfrak{m}^*$. Hence Γ_1 is a solvable subalgebra of $\mathfrak{m}$ containing $\mathfrak{h}+\Gamma$.

Now let β be any root of $\mathfrak{g}$. We claim that at least one of the two elements X_β or $X_{-\beta}$ is in Γ_1. In order to prove this we may clearly assume that $X_\beta \in \mathfrak{m}$ but $X_\beta \notin \Gamma_1$. Then $[H, X_\beta] = \beta(H)X_\beta$ $(H \in \mathfrak{h})$ and by going over to $\mathfrak{m}^*$ we get $[H^*, X_\beta^*] = \beta(H)X_\beta^*$ $(H \in \mathfrak{h})$. Since $X_\beta \notin \Gamma_1$, $X_\beta^* \neq 0$ and so we can choose a root γ^* of $\mathfrak{m}^*$ such that $\gamma^*(H^*) = \beta(H)$ for all $H \in \mathfrak{h}$. Define $\epsilon = 1$ or -1 according as γ^* is positive or negative and put $Y^* = X_{\epsilon\gamma^*}^* \in \Gamma_1^*$. Then if Y is any element in Γ_1 whose image in $\mathfrak{m}^*$ is Y^*, $Y \neq 0$ and

$$[H, Y] \equiv \epsilon\beta(H)Y \mod \Gamma \qquad (H \in \mathfrak{h}).$$

However Γ_1 is invariant and fully reducible under $ad\,\mathfrak{h}$. Therefore we can select an element $Z \in \Gamma_1$ $(Z \neq 0)$ such that

$$[H, Z] = \epsilon\beta(H)Z.$$

But then Z can differ from $X_{\epsilon\beta}$ only by a nonzero factor in C and therefore $X_{\epsilon\beta} \in \Gamma_1 \subset \mathfrak{m}$. This proves our assertion.

Put $B(X, Y) = sp(adX adY)$ $(X, Y \in \mathfrak{g})$. For any linear function λ on $\mathfrak{h}$ let H_λ' denote the (unique) element in $\mathfrak{h}$ such that $B(H_\lambda', H) = \lambda(H)$ for all $H \in \mathfrak{h}$. Define $\langle \lambda, \mu \rangle = B(H_\lambda', H_\mu') = \lambda(H_\mu')$ for any two linear functions λ, μ. Then it is well known [7] that if α is a root $H_\alpha' = c_\alpha H_\alpha$ where c_α is a *positive* real number. Let $\gamma_1, \cdots, \gamma_l$ be a fundamental system of roots of $\mathfrak{g}$. We introduce the lexicographic order (in $\mathfrak{F}_R$) corresponding to this system. Let Λ_0 be a linear function on $\mathfrak{h}$ such that $\Lambda_0(H_{\gamma_i})$ $1 \leqq i \leqq l$ are all positive integers. Now if α is a positive root $\alpha = d_1\gamma_1 + \cdots + d_l\gamma_l$ where d_i are nonnegative integers and therefore $H_\alpha' = d_1 H_{\gamma_1}' + \cdots + d_l H_{\gamma_l}'$. Hence

$$c_\alpha \Lambda_0(H_\alpha) = \Lambda_0(H_\alpha') = \sum_{i=1}^{l} d_i \Lambda_0(H_{\gamma_i}') = \sum_i d_i c_{\gamma_i} \Lambda_0(H_{\gamma_i})$$

and in view of the fact that $c_\alpha, c_{\gamma_i}, \Lambda_0(H_{\gamma_i})$ are all positive and not all d_i are zero we can conclude that $\langle \Lambda_0, \alpha \rangle$ and $\Lambda_0(H_\alpha)$ are both positive. Let π be an irreducible finite-dimensional representation of $\mathfrak{g}$ with the highest weight Λ_0. (We know from Theorem 1 of [5(b)] that such a representation exists). Since Γ_1 is a solvable algebra, by Lie's theorem we can choose an element $\psi_0 \neq 0$ in the representation space V and a linear function μ on Γ_1 such that $\pi(X)\psi_0 = \mu(X)\psi_0$ for all $X \in \Gamma_1$. Since $\mathfrak{h} \subset \Gamma_1$ it follows that μ coincides on $\mathfrak{h}$ with some weight Λ of π. Let β be a root such that $X_\beta \in \Gamma_1$.

Then ψ_0 is an eigenvalue of $\pi(X_\beta)$. But $\pi(X_\beta)$ and $\pi(X_{-\beta})$ are both nilpotent and therefore $\pi(X_\beta)\psi_0 = 0$ and $\pi(X_{-\beta}{}^r)\psi_0 = 0$ for r sufficiently large. Hence it follows from Lemma 1 that $\Lambda(H_\beta) \geqq 0$ and $\Lambda(H_\beta) \neq 0$ unless $\pi(X_{-\beta})\psi_0 = 0$. Let $(\beta_1, \cdots, \beta_s)$ be all the distinct roots β such that $X_\beta \in \Gamma_1$ and among these let $(\beta_1, \cdots, \beta_r)$ be all those for which $\Lambda(H_{\beta_i}) \neq 0$. Then $\Lambda(H_{\beta_i}) > 0$, $1 \leqq i \leqq r$ and in view of what we have proved above, $\mathfrak{g}$ is spanned by $\mathfrak{h}$, X_{β_i} and $X_{-\beta_i}$, $1 \leqq i \leqq s$. Hence it follows from Lemma 1 of [5(a)] that $\pi(\mathfrak{B})\psi_0 = V$ is spanned by elements of the form $\pi(X_{-\beta_1}{}^{m_1} \cdots X_{-\beta_s}{}^{m_s})\psi_0$ where $m_1, \cdots, m_s$ are nonnegative integers. Since $\pi(X_{-\beta_j})\psi_0 = 0$, $r < j \leqq s$, V is actually spanned by $\pi(X_{-\beta_1}{}^{m_1} \cdots X_{-\beta_r}{}^{m_r})\psi$. This proves that every weight of π must be of the form $\Lambda - (m_1\beta_1 + \cdots + m_r\beta_r)$, $(m_i \geqq 0)$. We now claim that Λ is an extreme weight of π. For otherwise let α be a root such that $\Lambda + \alpha$ and $\Lambda - \alpha$ are both weights. Then in view of the above result

$$\alpha = m_1\beta_1 + \cdots + m_r\beta_r, \qquad -\alpha = m_1'\beta_1 + \cdots + m_r'\beta_r$$

where $m_i, m_i' \geqq 0$. Hence $0 = (m_1 + m_1')\beta_1 + \cdots + (m_r + m_r')\beta_r$ and so

$$(m_1 + m_1')\langle \Lambda, \beta_1\rangle + (m_2 + m_2')\langle \Lambda, \beta_2\rangle + \cdots + (m_r + m_r')\langle \Lambda, \beta_r\rangle = 0.$$

But we have seen above that $\Lambda(H_{\beta_i}) > 0$ and therefore $\langle \Lambda, \beta_i\rangle > 0$, $i = 1, \cdots, r$. Hence the above equation implies that $m_i = m_i' = 0$, $1 \leqq i \leqq r$. This however is impossible since $\alpha \neq 0$. Therefore Λ is an extreme weight and we can conclude from Lemma 3 that $\Lambda = s\Lambda_0$ for some $s \in W$.

We recall that $\langle \Lambda_0, \alpha\rangle \neq 0$ for every root α. Hence $\langle \Lambda, \alpha\rangle = \langle \Lambda_0, s^{-1}\alpha\rangle \neq 0$ and therefore $\Lambda(H_\beta) > 0$ for every root β such that X_β lies in Γ_1. This shows in particular that for any root β exactly one of the two elements X_β and $X_{-\beta}$ is in Γ_1. Now put $\alpha_i = s\gamma_i$, $1 \leqq i \leqq l$. Then $(\alpha_1, \cdots, \alpha_l)$ is also a fundamental system of roots of $\mathfrak{g}$ and $\langle \Lambda, \alpha_i\rangle = \langle \Lambda_0, \gamma_i\rangle > 0$. Therefore $\Lambda(H_{\alpha_i}) > 0$ and hence $X_{\alpha_i} \in \Gamma_1 \subset \mathfrak{m}$. This proves the lemma.

COROLLARY 1. *Let P be a set of roots of $\mathfrak{g}$ (with respect to $\mathfrak{h}$) satisfying the following two conditions.*

(1) *For any root β exactly one of the two roots β and $-\beta$ lies in P;*

(2) *If α and β are in P and $\alpha + \beta$ is a root then $\alpha + \beta \in P$.*

Then it is possible to select roots $\alpha_i \in P$, $1 \leqq i \leqq l$ such that $(\alpha_1, \cdots, \alpha_l)$ is a fundamental system of roots and this system is unique (apart from a permutation of its elements). Moreover P is exactly the set of positive roots in the lexicographic order corresponding to $(\alpha_1, \cdots, \alpha_l)$.

It is clear that $\mathfrak{n} = \sum_{\beta \in P} CX_\beta$ is a subalgebra of $\mathfrak{g}$ and $\mathfrak{m} = \mathfrak{h} + \mathfrak{n}$ fulfils

all the conditions of Lemma 4. Hence we can choose a fundamental system $(\alpha_1, \cdots, \alpha_l)$ such that $X_{\alpha_i} \in \mathfrak{m}$. It is obvious that $\alpha_i \in P$. Let β be any positive root in the lexicographic order corresponding to this system. Then X_β is contained in the subalgebra of $\mathfrak{g}$ generated by X_{α_i}, $1 \leqq i \leqq l$ (see Lemma 18 of [5(b)]). Hence it follows from the second condition of our hypothesis that $X_\beta \in \mathfrak{n}$ and therefore $\beta \in P$. So every positive root is in P and hence in view of condition (1) no negative root can lie in P. This proves that P is exactly the set of all positive roots. Finally it is obvious that if α is a root in P, it lies in the fundamental system $(\alpha_1, \cdots, \alpha_l)$ if and only if it cannot be written in the form $\alpha = \beta_1 + \cdots + \beta_r$ where $\beta_1, \cdots, \beta_r$ are all roots in P and $r \geqq 2$. This shows that the system $(\alpha_1, \cdots, \alpha_l)$ is unique apart from order.

COROLLARY 2. *Let P be the set of all positive roots under any given order in $\mathfrak{F}_R$. Then there exists a fundamental system $(\alpha_1, \cdots, \alpha_l)$ consisting of positive roots and this system is unique apart from a permutation. Moreover P is also the set of all positive roots in the lexicographic order corresponding to $(\alpha_1, \cdots, \alpha_l)$.*

For it is obvious that P satisfies the two conditions of Corollary 1.

3. Representations with an extreme vector.

DEFINITION. *Let π be a representation of $\mathfrak{B}$ on a vector space V. An element ψ in V will be called an extreme vector of π (with respect to $\mathfrak{h}$) if the following two conditions are fulfilled:*

(1) *$\psi \neq 0$ and there exists a linear function Λ on $\mathfrak{h}$ such that $\pi(H)\psi = \Lambda(H)\psi$ for all $H \in \mathfrak{h}$.*

(2) *For any root β at least one of the two elements $\pi(X_\beta)\psi$, $\pi(X_{-\beta})\psi$ is zero.*

Suppose ψ is an extreme vector of π. Let $\mathfrak{m}$ be the subalgebra of $\mathfrak{g}$ consisting of all elements $X \in \mathfrak{g}$ such that ψ is an eigenvector of $\pi(X)$. Then $\mathfrak{m}$ satisfies the conditions of Lemma 4. If β is a root such that $\pi(X_\beta)\psi \neq 0$, $\pi(X_\beta)\psi$ belongs to the weight $\Lambda + \beta$ and therefore $X_\beta \notin \mathfrak{m}$. Hence if $X_\beta \in \mathfrak{m}$, $\pi(X_\beta)\psi = 0$. Therefore from Lemma 4 it is possible to choose a fundamental system of roots $(\alpha_1, \cdots, \alpha_l)$ such that $\pi(X_{\alpha_i})\psi = 0$, $i = 1, \cdots, l$. Then if we consider the lexicographic order corresponding to this system and set $\mathfrak{n} = \sum_{\beta > 0} CX_\beta$, $\mathfrak{n} \subset \mathfrak{m}$ (Lemma 18 of [5(b)] and therefore $\pi(X_\beta)\psi = 0$ for all $\beta > 0$. If $\beta_1, \cdots, \beta_r$ are all the distinct positive roots and we assume that $V = \pi(\mathfrak{B})\psi$, it follows from Lemma 2 that every weight

of π is of the form $\Lambda - (m_1\beta_1 + \cdots + m_r\beta_r)$, $(m_i \geqq 0)$ and therefore also of the form $\Lambda - (d_1\alpha_1 + \cdots + d_l\alpha_l)$ where $d_1, \cdots, d_l$ are nonnegative integers.

Conversely suppose we are already given an order in $\mathfrak{H}_R$ and ψ is an extreme vector of the representation π. We say that ψ is a *positive* extreme vector if $\pi(X_\beta)\psi = 0$ for all positive roots β. The above remarks show that any extreme vector is positive with respect to the lexicographic order corresponding to some fundamental system of roots.

LEMMA 5. *Let Λ be a linear function on $\mathfrak{h}$. Then there exists at most one maximal left ideal $\mathfrak{M}$ in $\mathfrak{B}$ such that*[4]

$$\mathfrak{M} \supset \sum_{\beta>0} \mathfrak{B}X_\beta + \sum_{\beta>0} \mathfrak{B}(H_\beta - \Lambda(H_\beta))$$

where β runs over all positive roots.

This is proved in exactly the same way as Lemma 1 of [5(b)].

COROLLARY. *Let π_1, π_2 be irreducible representations of $\mathfrak{B}$ on the vector spaces V_1, V_2 respectively. Suppose there exists an order in $\mathfrak{H}_R$ such that with respect to this order both π_1 and π_2 have positive extreme vectors belonging to the same weight Λ. Then π_1, π_2 are equivalent.*

Let ψ_i be a positive vector of π_i belonging to the weight Λ $(i = 1, 2)$. Let $\mathfrak{M}_i$ be the set of all $b \in \mathfrak{B}$ such that $\pi_i(b)\psi_i = 0$. Since π_i is irreducible $\mathfrak{M}_i$ is a maximal left ideal in $\mathfrak{B}$ and π_i is equivalent to the natural representation (see [5(c), § 3]) of $\mathfrak{B}$ on $\mathfrak{B}/\mathfrak{M}_i$. But $\mathfrak{M}_1 = \mathfrak{M}_2$ from Lemma 5 and so our assertion follows.

For reasons which we have explained briefly in Section 1 (see also [5(f)]), we shall now consider the following special case. Let $\mathfrak{g}_0$ be a semisimple Lie algebra over R. Henceforward I shall use the notation of my two previous papers [5(c), (d)]. Define $\mathfrak{k}_0$ and $\mathfrak{p}_0$ as in [5(c), § 2] and let $\mathfrak{h}_0$ be a maximal abelian subalgebra of $\mathfrak{k}_0$. *We shall assume that $\mathfrak{h}_0$ is also maximal abelian in $\mathfrak{g}_0$.* Complexify $\mathfrak{g}_0$ to $\mathfrak{g}$ and let $\mathfrak{h}$, $\mathfrak{p}$, $\mathfrak{k}$ denote the subspaces of $\mathfrak{g}$ spanned over C by $\mathfrak{h}_0$, $\mathfrak{p}_0$, $\mathfrak{k}_0$ respectively. Then $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$ and $[\mathfrak{h}, \mathfrak{k}] \subset \mathfrak{k}$, $[\mathfrak{h}, \mathfrak{p}] \subset \mathfrak{p}$. Therefore if β is any root (of $\mathfrak{g}$ with respect to $\mathfrak{h}$) it is clear that X_β lies either in $\mathfrak{k}$ or in $\mathfrak{p}$. We say that β is compact or noncompact according as $X_\alpha \in \mathfrak{k}$ or $X_\alpha \in \mathfrak{p}$. Two roots α and β are unlike if one of them is compact while the other is not. Otherwise they are like. Since $[\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}$, $[\mathfrak{k}, \mathfrak{p}] \subset \mathfrak{p}$ and $[X_\alpha, X_{-\alpha}] = H_\alpha \in \mathfrak{h} \subset \mathfrak{k}$, it follows that α and $-\alpha$ are always like. If α, β and $\alpha + \beta$ are roots, it is known that

[4] That $\mathfrak{M}$ always exists follows from Lemma 16

$[X_\alpha, X_\beta] = cX_{\alpha+\beta}$ where $c \neq 0$ $(c \in C)$. Therefore $\alpha+\beta$ is compact or not according as α, β are like or unlike.

Let π be a representation of $\mathfrak{g}$ on a vector space V with an extreme vector ψ_0 (with respect to $\mathfrak{h}$) belonging to the weight Λ_0. Introduce a lexicographic order so that $\pi(X_\beta)\psi_0 = 0$ for every positive root β. Let P be the set of all positive roots and P_k and P_n respectively the subsets of P consisting of all compact and noncompact roots in P. Then it is clear that

$$\mathfrak{k} = \mathfrak{h} + \sum_{\alpha \in P_k} (CX_\alpha + CX_{-\alpha}), \qquad \mathfrak{p} = \sum_{\beta \in P_n} (CX_\beta + CX_{-\beta}).$$

In accordance with our earlier notation [5(c)] we denote by $\mathfrak{X}$ the subalgebra of $\mathfrak{B}$ generated by $(1, \mathfrak{k})$. Since we are ultimately interested in the irreducible representations of the simply connected group corresponding to the algebra $\mathfrak{g}_0$, *we shall now make the further assumption that* $\dim \pi(\mathfrak{X})\psi_0$ *is finite and* $V = \pi(\mathfrak{B})\psi_0$.

Since $\mathfrak{X}$ is the universal enveloping algebra of $\mathfrak{k}$, we can identify representations of $\mathfrak{k}$ with those of $\mathfrak{X}$. Let σ denote the representation of $\mathfrak{k}$ on $\mathfrak{g}$ given by $\sigma(X)Y = [X, Y]$ $(X \in \mathfrak{k}, Y \in \mathfrak{g})$.

DEFINITION. *A root* β *of* $\mathfrak{g}$ *will be called totally positive if there exists a subset* P_β *of* P *such that* $\sigma(\mathfrak{X})X_\beta = \sum_{\gamma \in P_\beta} CX_\gamma$.

It is clear that every totally positive root is positive. Later (in Section 4) we shall give another equivalent definition of totally positive roots (see the corollary to Lemma 12). Let P_+ be the set of all totally positive roots and Q the remaining set of positive roots. Our main object at present is to prove the following theorem.

THEOREM 1. *For every* $\beta \in Q$, $\Lambda_0(H_\beta)$ *is a nonnegative integer and it is the least integer* $r \geqq 0$ *such that* $\pi(X_{-\beta}{}^{r+1})\psi_0 = 0$.

Let Q' denote the set of all positive roots β such that $\pi(X_{-\beta}{}^r)\psi_0 = 0$ for some $r \geqq 1$. Since $\pi(X_\beta)\psi_0 = 0$ for $\beta \in P$, it follows from Lemma 1 that in order to prove the above theorem it is enough to show that $Q \subset Q'$. For this we need some lemmas.

LEMMA 6. *Suppose* $\beta \in Q'$. *Then for any* $\psi \in V$ *we can choose an integer* $r \geqq 1$ *such that* $\pi(X_\beta{}^r)\psi = \pi(X_{-\beta}{}^r)\psi = 0$.

Let $\beta_1, \cdots, \beta_t$ be all the distinct positive roots. As we have already observed earlier $V = \pi(\mathfrak{B})\psi_0$ is then spanned by elements of the form $\pi(X_{-\beta_1}{}^{m_1} \cdots X_{-\beta_t}{}^{m_t})\psi_0$, $(m_j \geqq 0)$. Hence it is enough to prove our statement for $\psi = \pi(x)\psi_0$ where $x = X_{-\beta_1}{}^{m_1} \cdots X_{-\beta_t}{}^{m_t} \in \mathfrak{B}$. Now if $X \in \mathfrak{g}$,

$Xx = xX + [X, x]$. Therefore if we denote by adX the endomorphism $b \to [X, b]$ $(b \in \mathfrak{B})$ of $\mathfrak{B}$, it follows that (see pp. 40-41 of [5(b)])

$$X^p x = \sum_{0 \leqq q \leqq p} C_q{}^p \{(adX)^{p-q} x\} X^q$$

for any integer $p \geqq 0$. (Here $C_q{}^p = p!/(p-q)!q!$). However if $X = X_\beta$ or $X_{-\beta}$ the restriction of adX on $\mathfrak{g}$ is nilpotent and from this it follows easily that $(adX)x^m = 0$ if m is sufficiently large. Since $\beta \in Q'$, $\pi(X_\beta)\psi_0 = 0$ and we can choose an integer $s \geqq 1$ such that $\pi(X_{-\beta}{}^s)\psi_0 = 0$. Then if we put $p = m + s$ we get

$$X^{m+s} = \sum_{0 \leqq q \leqq m+s} C_q{}^{m+s} \{(adX) x^{m+s-q}\} X^q = \sum_{s < q \leqq m+s} C_q{}^{m+s} \{(adX) x^{m+s-q}\} X^q$$

since $(adX)x^m = 0$. On the other hand $\pi(X^q)\psi_0 = 0$ for $q \geqq s$ and therefore $\pi(X^{m+s})\psi = \pi(Xx^{m+s})\psi_0 = 0$. This proves our assertion.

COROLLARY. *Let W' be the subgroup of the Weyl group W generated by the Weyl reflexions s_β corresponding to $\beta \in Q'$. Then if Λ is a weight of π and $s \in W'$, $s\Lambda$ is also a weight of π.*

Suppose $s = s_{\beta_1} s_{\beta_2} \cdots s_{\beta_r}$ where $\beta_i \in Q'$. If $r = 1$ the result follows immediately from Lemmas 6 and 1. Hence assume $r \geqq 2$ and use induction on r. Then if $s' = s_{\beta_2} \cdots s_{\beta_r}$ it follows from our induction hypothesis that $s'\Lambda$ is a weight of π. Hence again by the above argument $s_{\beta_1}(s'\Lambda) = s\Lambda$ is also a weight of π.

LEMMA 7. *Let β be a positive root. Suppose $s\beta < 0$ for some $s \in W'$. Then $\beta \in Q'$.*

Let $\beta_1, \cdots, \beta_l$ be a fundamental system of positive roots (see Corollary 2 to Lemma 4). Then we know from Lemma 2 that every weight of π must be of the form $\Lambda_0 - (m_1\beta_1 + \cdots + m_l\beta_l)$ where m_i are nonnegative integers. Now suppose $\pi(X_{-\beta}{}^r)\psi_0 \neq 0$ for some positive integer. Then $\Lambda_0 - r\beta$ is a weight of π. Since $s \in W'$ it follows from the above corollary that both $s\Lambda_0$ and $s(\Lambda_0 - r\beta)$ are weights of π. Therefore

$$s\Lambda_0 = \Lambda_0 - (d_1\beta_1 + \cdots + d_l\beta_l) \qquad (d_i \geqq 0)$$

$$s\Lambda_0 - rs\beta = \Lambda_0 - (m_1\beta_1 + \cdots + m_l\beta_l) \qquad (m_i \geqq 0).$$

But $s\beta$ is a negative root and therefore $s\beta = -(e_1\beta_1 + \cdots + e_l\beta_l)$ $(e_i \geqq 0)$, e_i being integers which are not all zero. Hence

$$-r(e_1\beta_1 + \cdots + e_l\beta_l = rs\beta = s\Lambda_0 - (s\Lambda_0 - rs\beta)$$
$$= (m_1 - d_1)\beta_1 + \cdots + (m_l - d_l)\beta_l$$

and therefore $m_i = d_i - re_i \geqq 0$. Put $d = \max_i d_i$. Then $d \geqq \max_i re_i \geqq r$. This proves that $\pi(X_{-\beta}^{d+1})\psi_0 = 0$ and so $\beta \in Q'$.

Notice that if α is a positive compact root $\alpha \in Q'$. For the dimension of $\pi(\mathfrak{X})\psi_0$ is finite and the nonzero vectors among $\pi(X_{-\alpha}^r)\psi_0$ $(r \geqq 0)$ are linearly independent since they belong to the distinct weights $\Lambda_0 - r\alpha$. Hence $\pi(X_{-\alpha}^r)\psi_0 = 0$ for r sufficiently large. Therefore if W_k is the subgroup of W generated by the Weyl reflexions corresponding to the compact roots, $W_k \subset W'$.

Let $\mathfrak{c}_0$ be the center of $\mathfrak{k}_0$ and let $\mathfrak{k}_0' = [\mathfrak{k}_0, \mathfrak{k}_0]$. Then $\mathfrak{k}_0$ is the direct sum of $\mathfrak{c}_0$ and $\mathfrak{k}_0'$ (see [5(c), § 2]) and since $\mathfrak{h}_0$ is maximal abelian in $\mathfrak{k}_0$, $\mathfrak{c}_0 \subset \mathfrak{h}_0$. Let $\mathfrak{c}$ and $\mathfrak{k}'$ be the respective complexifications of $\mathfrak{c}_0$ and $\mathfrak{k}_0'$ in $\mathfrak{k}$. Put $\mathfrak{h}' = \mathfrak{h} \cap \mathfrak{k}'$.

LEMMA 8. *Let β be a positive root. If β is zero on $\mathfrak{c}$, $s\beta < 0$ for some $s \in W_k$ and therefore $\beta \in Q'$.*

It is obvious that every compact root is identically zero on $\mathfrak{c}$. Since $\beta - s\beta$ $(s \in W_k)$ is a linear combination of compact roots it follows that if β vanishes on $\mathfrak{c}$ the same holds for $s\beta$. Now consider $\lambda = \sum_{s \in W_k} s\beta$. Then $s_\alpha \lambda = \lambda$ and therefore $\lambda(H_\alpha) = 0$ for every compact root α. Since $\mathfrak{h}'$ is a Cartan subalgebra of the semisimple algebra $\mathfrak{k}'$ (see [5(c), Lemma 3]) it follows that the elements H_α corresponding to all compact roots α span $\mathfrak{h}'$. Hence $\lambda = 0$ on $\mathfrak{h}'$. But as we have seen $s\beta$ $(s \in W_k)$ and therefore also λ vanish on $\mathfrak{c}$. Hence $\lambda = 0$ and therefore $s\beta$ cannot all be positive. This proves that $s\beta < 0$ for some $s \in W_k$. But $W_k \subset W'$ and so we can conclude from Lemma 7 that $\beta \in Q'$.

We define the representation σ of $\mathfrak{k}$ on $\mathfrak{g}$ as before.

LEMMA 9. *Let γ be a root of $\mathfrak{g}$. Then $\sigma(\mathfrak{X})X_\gamma$ is irreducible under $\mathfrak{k}$.*

Since $\mathfrak{k}$ is reductive in $\mathfrak{g}$ [5(c), Lemma 3] the representation σ is semisimple. Also it is obvious that $\mathfrak{g}$ is fully reducible under $\sigma(\mathfrak{h})$ and CX_γ is the space of all elements $Y \in \mathfrak{g}$ such that $\sigma(H)Y = \gamma(H)Y$ for every $H \in \mathfrak{h}$. Let $\mathfrak{q}$ be a subspace of $\sigma(\mathfrak{X})X_\gamma$ which is invariant under $\sigma(\mathfrak{k})$. Since σ is semisimple there exists a subspace $\mathfrak{q}'$ complementary to $\mathfrak{q}$ in $\sigma(\mathfrak{X})X_\gamma$ which is also invariant under $\sigma(\mathfrak{k})$. Since $X_\gamma \in \mathfrak{q} + \mathfrak{q}'$, we can choose an element $Y \neq 0$ either in $\mathfrak{q}$ or $\mathfrak{q}'$ such that $\sigma(H)Y = \gamma(H)Y$ for all $H \in \mathfrak{h}$. Let $\mathfrak{q}_1 = \mathfrak{q}$ or $\mathfrak{q}'$ according as Y is in $\mathfrak{q}$ or $\mathfrak{q}'$. Then as we have remarked above $CY = CX_\gamma$ and therefore $\sigma(\mathfrak{X})X_\gamma = \sigma(\mathfrak{X})Y \subset \mathfrak{q}_1$. This proves that $\mathfrak{q}_1 = \sigma(\mathfrak{X})X_\gamma$ and so $\sigma(\mathfrak{X})X_\gamma$ is irreducible.

COROLLARY. *If β is any root such that $X_\beta \in \sigma(\mathfrak{X})X_\gamma$ then $\beta - \gamma$ is identically zero on $\mathfrak{c}$.*

For if $H \in \mathfrak{c}$, $\sigma(H)$ lies in the center of $\sigma(\mathfrak{X})$. Hence in view of the irreducibility proved above, it follows from Schur's lemma that there exists a complex number c such that $\sigma(H)Y = cY$ for all $Y \in \sigma(\mathfrak{X})X_\gamma$. Hence in particular if $Y = X_\beta$, $cX_\beta = \sigma(H)X_\beta = \beta(H)X_\beta$. Therefore $c = \beta(H)$ and similarly $c = \gamma(H)$. Hence $\beta(H) - \gamma(H) = 0$ for every $H \in \mathfrak{c}$.

Now we come to the proof of Theorem 1. Let γ_0 be a positive root which is not in Q'. We have to show that γ_0 is totally positive. Since $P_k \subset Q'$, γ_0 is not compact and therefore $X_{\gamma_0} \in \mathfrak{p}$. Put $\mathfrak{p}_{\gamma_0} = \sigma(\mathfrak{X})X_{\gamma_0}$. Then $\mathfrak{p}_{\gamma_0} \subset \mathfrak{p}$ and since $\mathfrak{p}$ is fully reducible under $\sigma(\mathfrak{h})$, $\mathfrak{p}_{\gamma_0}$ can be spanned by the elements X_β corresponding to a certain number of noncompact roots β. Let β_0 and β_1 respectively be the highest and lowest among all those roots β for which $X_\beta \in \mathfrak{p}_{\gamma_0}$. Then $\beta_0 \geqq \gamma_0 > 0$ and in order to prove that γ_0 is totally positive it would be enough to show that $\beta_1 > 0$. So suppose this is false and β_1 is negative. Since $\mathfrak{k} = \mathfrak{c} + \mathfrak{k}'$, it is obvious from Schur's lemma and Lemma 9 that $\mathfrak{p}_{\gamma_0}$ is irreducible under $\sigma(\mathfrak{k}')$. Let τ denote the irreducible representation of $\mathfrak{k}'$ on $\mathfrak{p}_{\gamma_0}$ defined by σ. Now $\mathfrak{h}'$ is a Cartan subalgebra of the semisimple algebra $\mathfrak{k}'$ and

$$\mathfrak{k}' = \mathfrak{h}' + \sum_{\alpha \in P_k} (CX_\alpha + CX_{-\alpha}).$$

Hence if we identify linear functions on $\mathfrak{h}'$ with those linear functions on $\mathfrak{h}$ which vanish on $\mathfrak{c}$, the roots of $\mathfrak{k}'$ (with respect to $\mathfrak{h}'$) coincide with the compact roots of $\mathfrak{g}$. For any root β of $\mathfrak{g}$ let β' denote the restriction of β on $\mathfrak{h}'$. It is clear that $[X_\alpha, X_{\beta_0}] = [X_{-\alpha}, X_{\beta_1}] = 0$ for all positive compact roots α. Therefore it follows from Lemma 2 that β_0' is the highest and β_1' the lowest weight (with respect to $\mathfrak{h}'$) of the irreducible representation τ of $\mathfrak{k}'$. Hence from Lemma 3, there exists an element t in the Weyl group of $\mathfrak{k}'$ with respect to $\mathfrak{h}'$, such that $\beta_1' = t\beta_0'$. This means that we can choose $s \in W_k$ such that $\beta_1 - s\beta_0$ is identically zero on $\mathfrak{h}'$. However $\beta_0 - s\beta_0$ is a linear combination of compact roots and therefore it vanishes on $\mathfrak{c}$. On the other hand from the above corollary $\beta_1 - \beta_0$ is also zero on $\mathfrak{c}$. Therefore $\beta_1 - s\beta_0$ vanishes on both $\mathfrak{h}'$ and $\mathfrak{c}$ and so $\beta_1 = s\beta_0$. Since $\beta_0 > 0$ and $s\beta_0 = \beta_1 < 0$, we conclude from Lemma 7 that $\beta_0 \in Q'$.

Now $X_{\gamma_0} \in \mathfrak{p}_{\gamma_0}$ and $\gamma_0 \notin Q'$. Choose the highest β such that $X_\beta \in \mathfrak{p}_{\gamma_0}$ and $\beta \notin Q'$. Then $\beta \geqq \gamma_0 > 0$ and $\beta \neq \beta_0$. We claim that $[X_\alpha, X_\beta] \neq 0$ for some positive compact root α. For otherwise suppose $[X_\alpha, X_\beta] = 0$ for every $\alpha \in P_k$. Let $\alpha_1, \cdots, \alpha_r$ be all the distinct positive compact roots of $\mathfrak{g}$. Since $\mathfrak{p}_{\gamma_0}$ is

irreducible under $\sigma(\mathfrak{k})$, it follows that $\sigma(\mathfrak{K})X_\beta = \mathfrak{p}_{\gamma_0}$. But then from Lemma 2 and the corollary to Lemma 9 we can conclude that

$$\beta_0 = \beta - (m_1\alpha_1 + \cdots + m_r\alpha_r) \qquad (m_i \geqq 0).$$

Since $\beta_0 \geqq \beta$ this implies that $m_1 = \cdots = m_r = 0$. However this is impossible as $\beta \neq \beta_0$. So let α be a positive compact root such that $[X_\alpha, X_\beta] \neq 0$. Then $\beta + \alpha$ is a root and $X_{\beta+\alpha} = c[X_\alpha, X_\beta] \in \mathfrak{p}_{\gamma_0}$ $(c \in C)$. Moreover $\beta + \alpha > \beta$ and therefore $\beta \in Q'$. Now $\beta + (\beta + \alpha)$ is not a root. For since $\beta + \alpha$ and β are both noncompact $\beta + (\beta + \alpha)$, if it is a root, must be compact. On the other hand $\beta - (\beta + \alpha) = -\alpha$ is also a compact root. Since $2\beta = \beta + (\beta + \alpha) - \alpha$, β is a linear combination of two compact roots and so it must be zero on $\mathfrak{c}$. Therefore from Lemma 8, $\beta \in Q'$. As this contradicts the definition of β, it follows that $\beta + (\beta + \alpha)$ is not a root while $\beta - (\beta + \alpha)$ is a root. Let ν be the largest integer $\geqq 0$ such that $\beta - \nu(\beta + \alpha)$ is a root. Then $\nu \geqq 1$ and

$$s_{\beta+\alpha}\beta = \beta - \nu(\beta + \alpha) \leqq \beta - (\beta + \alpha) = -\alpha < 0.$$

But $\beta + \alpha \in Q'$ and therefore from Lemma 7, $\beta \in Q'$ giving again a contradiction. This proves that $\beta_1 > 0$ and so γ_0 is totally positive.

4. Another characterization of totally positive roots. We shall now study the totally positive roots a little more closely. We call a root β totally negative if $-\beta$ is totally positive.

LEMMA 10. *Let β be a totally positive root. Then $s\beta$ $(s \in W_k)$ is also totally positive. Similarly if β is totally negative the same is true of $s\beta$.*

Let α be a compact root and let $\gamma = s_\alpha\beta$. Then it is known (see Weyl [7]) that X_γ can be written in the form $X_\gamma = c\sigma(Y^m)X_\beta$ where $c \in C$ and $Y = X_\alpha$ or $X_{-\alpha}$ and m is a nonnegative integer. Therefore from Lemma 9, $\sigma(\mathfrak{K})X_\gamma = \sigma(\mathfrak{K})X_\beta$. This shows that if β is totally positive the same holds for $s_\alpha\beta$ and therefore also for $s\beta$ $(s \in W_k)$. On the other hand if β is totally negative $s(-\beta) = -s\beta$ $(s \in W_k)$ is totally positive and therefore $s\beta$ is totally negative.

COROLLARY 1. *If β is totally positive it cannot be identically zero on $\mathfrak{c}$.*

Put $\lambda = \sum_{s \in W_k} s\beta$. Then $s\beta$ are all totally positive and hence positive. Therefore $\lambda > 0$. But $s_\alpha\lambda = \lambda$ for every compact root α and so λ is zero on $\mathfrak{h}'$. Hence λ cannot be identically zero on $\mathfrak{c}$. But $\beta - s\beta$ $(s \in W_k)$ being a linear combination of compact roots is zero on $\mathfrak{c}$. This shows that if w

is the order of the group W_k, λ coincides with $w\beta$ on $\mathfrak{c}$. Therefore β cannot vanish identically on $\mathfrak{c}$.

COROLLARY 2. *Every totally positive root is noncompact. If* $\mathfrak{c} = \{0\}$ *there are no totally positive roots.*

This is an immediate consequence of Corollary 1 above.

LEMMA 11. *Let* β *be a totally positive root and* γ *a noncompact root which is not totally negative. Then* $\beta + \gamma$ *is not a root.*

For otherwise suppose $\beta + \gamma$ is a root. Since β and γ are both noncompact $\alpha = \beta + \gamma$ must be compact. Then $-\gamma = \beta - \alpha$ and

$$X_{-\gamma} = c[X_{-\alpha}, X_\beta] \in \sigma(\mathfrak{X})X_\beta \qquad (c \in C).$$

Hence from Lemma 9, $\sigma(\mathfrak{X})X_{-\gamma} = \sigma(\mathfrak{X})X_\beta$ and therefore since β is totally positive the same must hold for $-\gamma$. But this means that γ is totally negative and so we get a contradiction.

Let θ denote the automorphism of $\mathfrak{g}$ of order 2 given by $\theta(X + Y) = X - Y$ $(X \in \mathfrak{k}, Y \in \mathfrak{p})$. We define a linear mapping $\tilde{\theta}$ of $\mathfrak{g}$ over R by the rule $\tilde{\theta}(cX) = \bar{c}\theta(X)$ $(c \in C, X \in \mathfrak{g}_0)$ where $\bar{c}$ is the complex conjugate of c. Then $\tilde{\theta}$ is the conjugation of $\mathfrak{g}$ with respect to its compact real form[5] $\mathfrak{u} = \mathfrak{k}_0 + (-1)^{\frac{1}{2}}\mathfrak{p}_0$ (see [5(c), § 2]) and it is known (see Weyl [7]) that one can choose the elements X_γ corresponding to the various roots γ in such a way that $X_\gamma - X_{-\gamma}$ and $(-1)^{\frac{1}{2}}(X_\gamma + X_{-\gamma})$ both lie in $\mathfrak{u}$ (without disturbing our earlier condition that $\gamma(H_\gamma) = 2$ for $H_\gamma = [X_\gamma, X_{-\gamma}]$). Assuming that this has been done, it follows immediately that $\tilde{\theta}(X_\gamma) = -\epsilon_\gamma X_{-\gamma}$ where $\epsilon_\gamma = 1$ or -1 according as γ is compact or not.

LEMMA 12. *Let* $\alpha_1, \cdots, \alpha_r$ *be all the distinct positive compact roots and* β *a totally positive root. Then if* γ *is any root of the form* $\beta + m_1\alpha_1 + \cdots + m_r\alpha_r$ *(where* $m_1, \cdots, m_r$ *are integers),* γ *is totally positive and* $X_\gamma \in \sigma(\mathfrak{X})X_\beta$.

We shall make use of an argument due to E. Cartan [3(c)]. Define the bilinear form $B(X, Y)$ on $\mathfrak{g}$ as before. Then $-B(X, \tilde{\theta}(X))$ $(X \in \mathfrak{g})$ is a positive definite Hermitian form on $\mathfrak{g}$. Since β is noncompact, $\mathfrak{p}_\beta = \sigma(\mathfrak{X})X_\beta$ is a subspace of $\mathfrak{p}$. Let $\mathfrak{q}$ denote the orthogonal complement[6]

[5] We fix once for all one square-root of -1 in C and denote it by $(-1)^{\frac{1}{2}}$.

[6] I take this opportunity to correct a slip in one of my earlier papers [5(d), p. 29 bottom paragraph]. There the orthogonal complement is taken with respect to $B(X, Y)$ instead of $B(X, \tilde{\theta}, Y))$. This is a mistake which however is easy to rectify without making any substantial change in the argument.

of $\mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta)$ in $\mathfrak{p}$ (with respect to the above Hermitian form). Then we shall prove that $[\mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta), \mathfrak{q}] = \{0\}$. Since $\mathfrak{q} = \tilde{\theta}(\mathfrak{q})$, it is enough to show that $[\mathfrak{p}_\beta, \mathfrak{q}] = \{0\}$. Let $X \in \mathfrak{p}_\beta$, $Y \in \mathfrak{q}$ and $Z \in \mathfrak{k}$. Then

$$B([X, Y], \tilde{\theta}(Z)) = B(Y, [\tilde{\theta}(Z), X]) = 0$$

since $[\tilde{\theta}(Z), X]$ lies in $\mathfrak{p}_\beta$ and $\mathfrak{q}$ is orthogonal to $\tilde{\theta}(\mathfrak{p}_\beta)$. But $[X, Y] \in [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}$ and so this shows that $[X, Y] = 0$ and our assertion is proved. Let $\mathfrak{k}_\beta$ be the centralizer of $\mathfrak{q}$ in $\mathfrak{k}$. Put $\mathfrak{g}_\beta = \mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta) + \mathfrak{k}_\beta$. We claim that $\mathfrak{g}_\beta$ is an ideal in $\mathfrak{g}$. Since $\mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta)$ is invariant under $\sigma(\mathfrak{k})$, the same holds for $\mathfrak{q}$ and $\mathfrak{k}_\beta$. Therefore $[\mathfrak{k}, \mathfrak{g}_\beta] \subset \mathfrak{g}_\beta$. Moreover $[\mathfrak{q}, \mathfrak{g}_\beta] = \{0\}$. Since $\mathfrak{g} = \mathfrak{k} + \mathfrak{q} + \mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta)$, it is enough to prove that $[\mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta), \mathfrak{g}_\beta] \subset \mathfrak{g}_\beta$. But $\mathfrak{g}_\beta = \mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta) + \mathfrak{k}_\beta$ and $\mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta)$ is invariant under $\sigma(\mathfrak{k})$. So it would be sufficient to show that if X, Y are in $\mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta)$, $[X, Y]$ lies in $\mathfrak{k}_\beta$. However $[X, Y] \in \mathfrak{k}$ and if $Z \in \mathfrak{q}$

$$[[X, Y], Z] = [[X, Z], Y] + [X, [Y, Z]] = 0$$

since $[X, Z] = [Y, Z] = 0$. This proves that $[X, Y] \in \mathfrak{k}_\beta$ and therefore $\mathfrak{g}_\beta$ is an ideal in $\mathfrak{g}$. Let $\mathfrak{l}$ be the orthogonal complement of $\mathfrak{k}_\beta$ in $\mathfrak{k}$. Then $\mathfrak{g}' = \mathfrak{q} + \mathfrak{l}$ is clearly the orthogonal complement of $\mathfrak{g}_\beta$ in $\mathfrak{g}$. Let $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}'$. Then if $Z \in \mathfrak{g}_\beta$,

$$B([X, Y], \tilde{\theta}(Z)) = -B(Y, [X, \tilde{\theta}(Z)]).$$

But

$$\tilde{\theta}([X, \tilde{\theta}(Z)]) = [\tilde{\theta}(X), Z] \in \mathfrak{g}_\beta$$

and since $\mathfrak{g}'$ is orthogonal to $\mathfrak{g}_\beta$, $B(Y, [X, \tilde{\theta}(Z)]) = 0$. Hence $[X, Y]$ being orthogonal to $\mathfrak{g}_\beta$ lies in $\mathfrak{g}'$. This shows that $\mathfrak{g}'$ is also an ideal in $\mathfrak{g}$. Hence $[\mathfrak{g}', \mathfrak{g}_\beta] \subset \mathfrak{g}' \cap \mathfrak{g}_\beta = \{0\}$. Moreover since $\mathfrak{q} = \tilde{\theta}(\mathfrak{q})$ it follows that $\mathfrak{k}_\beta = \tilde{\theta}(\mathfrak{k}_\beta)$, $\tilde{\theta}(\mathfrak{g}_\beta) = \mathfrak{g}_\beta$ and $\tilde{\theta}(\mathfrak{g}') = \mathfrak{g}'$. Now suppose $X \in \mathfrak{c}$. Since $\mathfrak{k} = \mathfrak{k}_\beta + \mathfrak{l}$, $X = Y + Z$ where $Y \in \mathfrak{k}_\beta$ and $Z \in \mathfrak{l}$. But $[\mathfrak{k}_\beta, \mathfrak{l}] = \{0\}$ and so it is obvious that both Y and Z are in $\mathfrak{c}$. This shows that $\mathfrak{c} = \mathfrak{c} \cap \mathfrak{k}_\beta + \mathfrak{c} \cap \mathfrak{l}$.

Now suppose $\gamma = \beta + m_1\alpha_1 + \cdots + m_r\alpha_r$ is a root. Then γ and β coincide on $\mathfrak{c}$. Since $X_\beta \in \mathfrak{g}_\beta$, it commutes with $\mathfrak{l}$ and therefore β vanishes identically on $\mathfrak{c} \cap \mathfrak{l}$. Hence the same holds for γ. Moreover since both $\mathfrak{g}_\beta$ and $\mathfrak{g}'$ are invariant under $\sigma(\mathfrak{h})$ it is clear that X_γ lies either in $\mathfrak{g}_\beta$ or $\mathfrak{g}'$. We claim $X_\gamma \in \mathfrak{g}_\beta$. For otherwise if $X_\gamma \in \mathfrak{g}'$ it commutes with $\mathfrak{g}_\beta$ and therefore γ vanishes identically on $\mathfrak{c} \cap \mathfrak{k}_\beta$. But we have already seen that γ is zero on $\mathfrak{c} \cap \mathfrak{l}$. Therefore γ and hence also β vanishes identically on $\mathfrak{c}$. But since β is totally positive this is impossible in view of Corollary 1 to Lemma 10. This proves that $X_\gamma \in \mathfrak{g}_\beta$. Now γ cannot be a compact root since it co-

incides with β on $\mathfrak{c}$ and therefore is not identically zero on $\mathfrak{c}$. Hence $X_\gamma \in \mathfrak{p} \cap \mathfrak{g}_\beta = \mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta)$. Since $\mathfrak{p}_\beta$ and $\tilde{\theta}(\mathfrak{p}_\beta)$ are both invariant under $\sigma(\mathfrak{h})$ either $X_\gamma \in \mathfrak{p}_\beta$ or $X_\gamma \in \tilde{\theta}(\mathfrak{p}_\beta)$. But $\tilde{\theta}(X_\gamma) = -\epsilon_\gamma X_{-\gamma}$ and so either X_γ or $X_{-\gamma}$ lies in $\mathfrak{p}_\beta$. Now if $X_{-\gamma} \in \mathfrak{p}_\beta$, it follows from the corollary to Lemma 9 that $\beta + \gamma$ is zero on $\mathfrak{c}$. But we have seen above that $\beta - \gamma$ is zero on $\mathfrak{c}$. Therefore $2\beta = (\beta + \gamma) + (\beta - \gamma)$ is also zero on $\mathfrak{c}$. Since β is totally positive this is impossible (Corollary 1 to Lemma 10) and therefore $X_\gamma \in \mathfrak{p}_\beta$. Hence from Lemma 9 $\sigma(\mathfrak{X})X_\gamma = \sigma(\mathfrak{X})X_\beta$ and this proves that γ is totally positive.

COROLLARY. *Let $\alpha_1, \cdots, \alpha_r$ be all the positive compact roots of $\mathfrak{g}$. Then a root β is totally positive if and only if every root of the form $\beta + m_1\alpha_1 + \cdots + m_r\alpha_r$* (where $m_1, \cdots, m_r$ are integers) *is positive.*

Suppose $\gamma = \beta + m_1\alpha_1 + \cdots + m_r\alpha_r$ is a root. If β is totally positive, it follows from the above lemma that γ is positive. Conversely suppose every root γ of this form is positive. Then it is clear that $\beta \neq \pm \alpha_i$, $1 \leqq i \leqq r$ and therefore β is noncompact and $\sigma(\mathfrak{X})X_\beta$ is spanned by the elements X_γ corresponding to certain roots γ of the above form. Since every such γ is positive, it follows from our definition that β is totally positive. The above corollary thus gives an alternative definition of total positivity.

5. Some further results on totally positive roots. First we recall a few well-known facts in the theory of semisimple Lie algebras. Let $\mathfrak{a}$ be an ideal in $\mathfrak{g}$ and β a root of $\mathfrak{g}$. We say that β is a root of $\mathfrak{a}$ if $X_\beta \in \mathfrak{a}$. Let $(\beta_1, \cdots, \beta_l)$ be a fundamental system of roots of $\mathfrak{g}$ and suppose among these $\beta_1, \cdots, \beta_m$ are all the roots of $\mathfrak{a}$. Then every root of $\mathfrak{a}$ is a linear combination of $\beta_1, \cdots, \beta_m$. Let s_i denote the Weyl reflexion corresponding to β_i and let $W_\mathfrak{a}$ be the subgroup of the Weyl group W generated by $s_1, \cdots, s_m$. Then if β is any root of $\mathfrak{a}$ the corresponding Weyl reflexion s_β is contained in $W_\mathfrak{a}$.

Now as before let P_+ be the set of all totally positive roots. Put $\mathfrak{p}_+ = \sum_{\beta \in P_+} CX_\beta$ and $\mathfrak{p}_- = \sum_{\beta \in P_+} CX_{-\beta}$ so that $\mathfrak{p}_- = \tilde{\theta}(\mathfrak{p}_+)$. Let $\mathfrak{q}$ be the orthogonal complement of $\mathfrak{p}_+ + \mathfrak{p}_-$ in $\mathfrak{p}$ (under the Hermitian form $B(X, \tilde{\theta}(X))$). Then if $\mathfrak{l}$ is the centralizer of $\mathfrak{q}$ in $\mathfrak{k}$ we can show, as in the proof of Lemma 12, that $\mathfrak{g}_+ = \mathfrak{p}_+ + \mathfrak{p}_- + \mathfrak{l}$ is an ideal in $\mathfrak{g}$. Let $\mathfrak{g}'$ be orthogonal complement of $\mathfrak{g}_+$ in $\mathfrak{g}$. Then $\mathfrak{g}'$ is also an ideal in $\mathfrak{g}$ and $[\mathfrak{g}', \mathfrak{g}_+] = \{0\}$. It is obvious that $\mathfrak{g}' \cap \mathfrak{k}$ and $\mathfrak{g}' \cap \mathfrak{p}$ are the orthogonal complements of $\mathfrak{l}$ and $\mathfrak{p}_+ + \mathfrak{p}_-$ in $\mathfrak{k}$ and $\mathfrak{p}$ respectively and $\mathfrak{g}' = \mathfrak{g}' \cap \mathfrak{k} + \mathfrak{g}' \cap \mathfrak{p}$. Now if β is a root, X_β lies

either in $\mathfrak{g}_+$ or $\mathfrak{g}'$. Since $\mathfrak{g}_+ \cap \mathfrak{p} = \mathfrak{p}_+ + \mathfrak{p}_-$, it follows that if β is a noncompact root of $\mathfrak{g}_+$ it is either totally positive or totally negative.

Let $(\beta_1, \cdots, \beta_l)$ be a fundamental system of positive roots of $\mathfrak{g}$ (Corollary 2 to Lemma 4). If γ is a positive root which is not contained in this fundamental system, then it is well known that $\gamma - \beta_i$ must also be a positive root for some i $(1 \leqq i \leqq l)$. On the other hand $\beta_i - \beta$ can never be a positive root for any root $\beta > 0$. Hence a positive root is contained in the set $(\beta_1, \cdots, \beta_l)$ if and only if $\gamma - \beta$ is never a positive root for any root $\beta > 0$.

Now it follows from the definition of total positivity that $\mathfrak{p}_+$ is invariant under $\sigma(\mathfrak{k})$. Hence $\mathfrak{p}_+ = \sum_{\nu=1}^{q} \mathfrak{q}_\nu$ where $\mathfrak{q}_\nu$, $1 \leqq \nu \leqq q$, are nonzero subspaces of $\mathfrak{p}_+$ which are invariant and irreducible under $\sigma(\mathfrak{k})$ and the sum is direct. (If $\mathfrak{p}_+ = \{0\}$, $q = 0$). Let γ_ν be the lowest root such that $X_{\gamma_\nu} \in \mathfrak{q}_\nu$ and let α be a positive compact root. We claim $\gamma_\nu - \alpha$ is not a root. For otherwise it follows from Lemma 12 that $X_{\gamma_\nu - \alpha} \in \mathfrak{q}_\nu$ which however is impossible since $\gamma_\nu - \alpha < \gamma_\nu$. Now let β be a noncompact positive root and suppose $\alpha = \gamma_\nu - \beta$ is a root. Since γ_ν and β are both noncompact, α is compact and $\beta = \gamma_\nu - \alpha$. Hence by the above argument α cannot be positive. This proves that $\gamma_\nu - \beta$ can never be a positive root for any root $\beta > 0$. Therefore it follows from our criterion above that γ_ν is a root of the fundamental system $(\beta_1, \cdots, \beta_l)$.

Now suppose $(\beta_1, \cdots, \beta_r)$ are roots of $\mathfrak{g}_+$ while $(\beta_{r+1}, \cdots, \beta_l)$ are roots of $\mathfrak{g}'$ $(0 \leqq r \leqq l)$. Then if α is a positive compact root of $\mathfrak{g}_+$,

$$\alpha = m_1\beta_1 + \cdots + m_r\beta_r \qquad (m_i \geqq 0).$$

We recall that if γ is any compact root $\sum_{s \in W_k} s\gamma = 0$. Now suppose $\beta_1, \cdots, \beta_t$ are compact while $\beta_{t+1}, \cdots, \beta_r$ are noncompact (and hence totally positive). Then

$$0 = \sum_{s \in W_k} s\alpha = \sum_{i=1}^{r} m_i \sum_{s \in W_k} s\beta_i = \sum_{t < i \leqq r} m_i \sum_{s \in W_k} s\beta_i$$

But from Lemma 10 $s\beta_i$ $(i > t)$ is also totally positive. Since $m_i \geqq 0$ the above sum cannot be zero unless $m_i = 0$ $(t < i \leqq r)$. Therefore $\alpha = m_1\beta_1 + \cdots + m_t\beta_t$. On the other hand let γ be any positive noncompact root of $\mathfrak{g}_+$. Then $X_\gamma \in \mathfrak{q}_\nu = \sigma(\mathfrak{X}) X_{\gamma_\nu}$ for some ν. Since $[\mathfrak{g}', \mathfrak{g}_+] = \{0\}$, it is clear that $\mathfrak{q}_\nu$ is irreducible under $\sigma(\mathfrak{l})$ and therefore $\gamma - \gamma_\nu$ is a linear combination of compact roots of $\mathfrak{g}_+$. Hence $\gamma = \gamma_\nu + m_1\beta_1 + \cdots + m_t\beta_t$ where $m_1, \cdots, m_t$ are real numbers. But γ_ν is contained in the set $(\beta_1, \cdots, \beta_r)$ and since it is noncompact, $\gamma_\nu \neq \beta_i$, $1 \leqq i \leqq t$. Therefore

since $\gamma > 0$ and $(\beta_1, \cdots, \beta_l)$ is a fundamental system, $m_1, \cdots, m_l$ must all be nonnegative integers. Thus we have obtained the following result.

LEMMA 13. *Let $(\beta_1, \cdots, \beta_l)$ be a fundamental system of positive roots of $\mathfrak{g}$. Let $(\alpha_1, \cdots, \alpha_r)$ be all the (distinct) compact roots of $\mathfrak{g}_+$ and $(\gamma_1, \cdots, \gamma_t)$ all the (distinct) noncompact roots of $\mathfrak{g}_+$ which occur in this fundamental system. Then every positive root γ of $\mathfrak{g}_+$ can be written in the form*

$$\gamma = \mu_1\gamma_1 + \cdots + \mu_t\gamma_t + m_1\alpha_1 + \cdots + m_r\alpha_r$$

where μ_i, m_j are all nonnegative integers. Moreover $\mu_1 + \cdots + \mu_t = 0$ or 1 according as γ is compact or not.

COROLLARY 1. *$t = \dim(\mathfrak{c} \cap \mathfrak{g}_+)$ and $r = \dim(\mathfrak{h}' \cap \mathfrak{g}_+)$.*

For it is easy to see that $\mathfrak{h} = \mathfrak{h} \cap \mathfrak{g}_+ + \mathfrak{h} \cap \mathfrak{g}'$ and $\mathfrak{c} = \mathfrak{c} \cap \mathfrak{g}_+ + \mathfrak{c} \cap \mathfrak{g}'$. From this it follows that $r + t = \dim(\mathfrak{h} \cap \mathfrak{g}_+)$ and $r = \dim(\mathfrak{h}' \cap \mathfrak{g}_+)$. Since $\mathfrak{h} \cap \mathfrak{g}_+ = \mathfrak{h}' \cap \mathfrak{g}_+ + \mathfrak{c} \cap \mathfrak{g}_+$, $t = \dim(\mathfrak{c} \cap \mathfrak{g}_+)$.

COROLLARY 2. *If the algebra $\mathfrak{g}_0$ is simple, only the following two cases are possible.*

(1) *There are no totally positive roots.*

(2) *Every noncompact positive root is totally positive.*

Suppose we are in case (2) and $\mathfrak{g}_0 \neq \mathfrak{k}_0$. Then there is exactly one noncompact root in the fundamental system $(\beta_1, \cdots, \beta_l)$ and $\dim_R \mathfrak{c}_0 = 1$.

Let η denote the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$ so that

$$\eta(X + (-1)^{\frac{1}{2}}Y) = X - (-1)^{\frac{1}{2}}Y \qquad (X, Y \in \mathfrak{g}_0).$$

Then $\eta = \tilde{\theta} \circ \theta$ and since $\mathfrak{g}_+$ and $\mathfrak{g}'$ are invariant under both $\tilde{\theta}$ and θ they are also invariant under η. Therefore $\mathfrak{g}_0 = \mathfrak{g}_0 \cap \mathfrak{g}_+ + \mathfrak{g}_0 \cap \mathfrak{g}'$ and $\mathfrak{g}_+$, $\mathfrak{g}'$ are spanned over C by $\mathfrak{g}_0 \cap \mathfrak{g}_+$ and $\mathfrak{g}_0 \cap \mathfrak{g}'$ respectively. Hence it follows that if $\mathfrak{g}_0$ is simple either $\mathfrak{g}_+$ or $\mathfrak{g}'$ is $\{0\}$ and so we get the above two cases. Now suppose we are in case (2). Then every positive root is either compact or totally positive. Therefore if $\mathfrak{g}_0 \neq \mathfrak{k}_0$, totally positive roots must exist and so $t \geqq 1$ in the notation of Lemma 13. On the other hand it follows from the simplicity of $\mathfrak{g}_0$ (see Cartan [3(e), p. 247]) that $\dim_R \mathfrak{c}_0 \leqq 1$. Therefore from Corollary 1 above $t = \dim_R \mathfrak{c}_0 = 1$.

LEMMA 14. *Let $(\beta_1, \cdots, \beta_q)$ be all those roots in the fundamental system $(\beta_1, \cdots, \beta_l)$ (of Lemma 13) which are not totally positive and let W' be the subgroup of the Weyl group W generated by the Weyl reflexions*

corresponding to β_i, $1 \leqq i \leqq q$. *Then if* β *is any positive root of* $\mathfrak{g}$ *which is not totally positive we can find an index* i $(1 \leqq i \leqq q)$ *and an element* $s \in W'$ *such that* $\beta = s\beta_i$.

For any root γ let s_γ denote the corresponding Weyl reflexion. Put $s_i = s_{\beta_i}$, $1 \leqq i \leqq l$. Then it is well known that $s_1, \cdots, s_l$ generate W. Now if some β_i is a root of $\mathfrak{g}'$, it cannot be totally positive and therefore $i \leqq q$. So we may suppose that $(\beta_1, \cdots, \beta_d)$ $(0 \leqq d \leqq q)$ are all the roots of $\mathfrak{g}'$ among $(\beta_1, \cdots, \beta_l)$. Since $[\mathfrak{g}_+, \mathfrak{g}'] = \{0\}$, it is easy to see that if γ and γ' are roots of $\mathfrak{g}_+$ and $\mathfrak{g}'$ respectively then s_γ and $s_{\gamma'}$ commute and $s_\gamma \gamma' = \gamma'$, $s_{\gamma'}\gamma = \gamma$. Therefore in particular s_i and s_j commute if $i \leqq d$ and $j > d$. Let W'' and W_+ be the subgroups of W generated by s_i, $1 \leqq i \leqq d$ and s_j, $d < j \leqq l$ respectively and let β be a positive root which is not totally positive. First suppose β is a root of $\mathfrak{g}'$. Then $\beta = s\beta_i$ for some $s \in W$ and some i $(1 \leqq i \leqq l)$. But we can write $s^{-1} = uv$ where $u \in W''$ and $v \in W_+$. Then $v\beta = \beta$ and therefore $\beta_i = s^{-1}\beta = u\beta$. Since $u \in W''$, $u\beta - \beta$, and therefore also β_i, is a linear combination of roots of $\mathfrak{g}'$. This proves that β_i is a root of $\mathfrak{g}'$ and therefore $i \leqq d$. Since $u^{-1} \in W'' \subset W'$ and $\beta = u^{-1}\beta_i$, our assertion is true in this case. On the other hand if β is a root of $\mathfrak{g}_+$, it must be compact. Now $\mathfrak{k}' \cap \mathfrak{g}_+$ being an ideal in the semisimple algebra $\mathfrak{k}'$, is semisimple and the compact roots of $\mathfrak{g}_+$ may be identified in the obvious manner with the roots of $\mathfrak{k}' \cap \mathfrak{g}_+$ with respect to $\mathfrak{h}' \cap \mathfrak{g}_+$. Then it follows from Lemma 13 that $(\beta_{d+1}, \cdots, \beta_q)$ is a fundamental system of roots of $\mathfrak{k}' \cap \mathfrak{g}_+$ and therefore $\beta = s\beta_j$ where $d < j \leqq q$ and s lies in the group generated by $(s_{d+1}, \cdots, s_q)$. This completes the proof of the lemma.

We shall now prove a result which will be needed in a subsequent paper.

LEMMA 15. *Let* γ *be a totally positive root and* α *a compact root. Let* k *and* k' *respectively be the largest nonnegative integers such that* $\gamma - k\alpha$ *and* $\gamma + k'\alpha$ *are roots. Then* $k + k' \leqq 2$.

Put $\gamma_0 = \gamma + k'\alpha$. Then it is known (see Weyl [7]) that $s_\alpha \gamma_0 = \gamma - k\alpha$ and $\gamma_0 - \nu\alpha$ is a root for an integer ν if and only if $0 \leqq \nu \leqq k + k' = \gamma_0(H_\alpha)$. So we have to prove that $\gamma_0(H_\alpha) \leqq 2$. Hence let us suppose that $\gamma_0(H_\alpha) \geqq 3$. Then $\gamma_0 - \alpha$ is a root and therefore also $\alpha - \gamma_0$. Hence α and $\alpha - \gamma_0$ are roots while $\alpha + \gamma_0$ is not a root. This implies (see Weyl [7]) that $s_{\gamma_0}\alpha \leqq \alpha - \gamma_0$ or $\alpha(H_{\gamma_0}) \geqq 1$ and therefore $\gamma_0(H_\alpha)\alpha(H_{\gamma_0}) \geqq 3$. Now consider the root $\gamma' = s_{\gamma_0} s_\alpha \gamma_0$. It is easy to verify that $\gamma' = m\gamma_0 - \gamma_0(H_\alpha)\alpha$ where $m = \gamma_0(H_\alpha)\alpha(H_{\gamma_0}) - 1$. Since γ is totally positive the same holds for γ_0 and $s_\alpha \gamma_0$ (corollary to Lemma 12) and therefore $\gamma' = s_{\gamma_0} s_\alpha \gamma_0$ is a root of $\mathfrak{g}_+$.

However $m \geqq 2$ and so we get a contradiction with Lemma 13. This proves that $\gamma_0(H_a) \leqq 2$.

6. The converse of Theorem 1. Assuming that we are given an order in $\mathfrak{F}_R$ we shall now pove the following converse of Theorem 1.

THEOREM 2. *Let Λ_0 be a linear function on $\mathfrak{h}$ such that $\Lambda_0(H_\beta)$ is a nonnegative integer for every positive root β which is not totally positive. Then there exists a representation π of $\mathfrak{B}$ on a vector space V and a vector $\psi \neq 0$ in V such that the following condition holds:*

(1) *$\pi(X_\beta)\psi = 0$ for every positive root β.*

(2) *$\pi(H)\psi = \Lambda_0(H)\psi$ for all H in $\mathfrak{h}$.*

(3) *π is irreducible and $\dim \pi(\mathfrak{X})\psi$ is finite.*

Moreover such a representation is unique up to equivalence.

Choose a fundamental system $(\beta_1, \cdots, \beta_l)$ of positive roots and let $(\beta_1, \cdots, \beta_q)$ be all the roots in this system which are not totally positive. Then in the notation of Lemma 13, $(\gamma_1, \cdots, \gamma_t)$ are exactly the remaining roots of our fundamental system. Consider the left ideal

$$\mathfrak{M}_0 = \sum_{\beta \in P} \mathfrak{B} X_\beta + \sum_{\beta \in P} \mathfrak{B}(H_\beta - \Lambda_0(H_\beta)) + \sum_{1 \leqq i \leqq q} \mathfrak{B}\, Y_i^{\lambda_i+1}$$

in $\mathfrak{B}$. Here P is the set of all positive roots, $\lambda_i = \Lambda_0(H_{\beta_i})$ and $Y_i = X_{-\beta_i}$, $1 \leqq i \leqq q$. First we claim that $\mathfrak{M}_0 \neq \mathfrak{B}$ (cf. Lemma 8 of [5(b)]). For otherwise if $\mathfrak{M}_0 = \mathfrak{B}$ we can choose $z_i \in \mathfrak{B}$ $(1 \leqq i \leqq q)$ such that

$$1 \equiv \sum_{1 \leqq i \leqq q} z_i Y_i^{\lambda_i+1} \bmod \mathfrak{P}_{\Lambda_0}$$

where $\mathfrak{P}_{\Lambda_0} = \sum_{\beta \in P} \mathfrak{B} X_\beta + \sum_{\beta \in P} \mathfrak{B}(H_\beta - \Lambda\,(H_\beta))$. Since 1 is of rank zero (see Section 2 for the definition of rank) while $Y_i^{\lambda_i+1}$ is of rank $-(\lambda_i+1)\beta_i$, we may assume (see the proof of Lemma 8 of [5(b)]) that z_i is of rank $(\lambda_i+1)\beta_i$. Put $X_i = X_{\beta_i}$, $Y_i = X_{-\beta_i}$ and $H_i = H_{\beta_i}$ $(1 \leqq i \leqq l)$. Then $(1, X_i, H_i, Y_i)_{1 \leqq i \leqq l}$ generate $\mathfrak{B}$ (Lemma 18 of [5(b)]). Therefore if $\mathfrak{U}, \mathfrak{H}, \mathfrak{V}$ are the subalgebras of $\mathfrak{B}$ generated by $(1, X_1, \cdots, X_l)$, $(1, H_1, \cdots, H_l)$ and $(1, Y_1, \cdots, Y_l)$ respectively, it is easy to see that $\mathfrak{B} = \mathfrak{V}\mathfrak{H}\mathfrak{U}$ (see the proof of Lemma 2 of [5(b)]). However in view of the argument given in the proof of Lemma 8 of [5(b), p. 37] we may assume that $z_i \in \mathfrak{U}$. Since the product $X_{i_1} X_{i_2} \cdots X_{i_r}$ is of rank $\beta_{i_1} + \beta_{i_2} + \cdots \beta_{i_r}$ and since $\beta_1, \cdots, \beta_l$ are linearly independent, it follows that the only elements in $\mathfrak{U}$ of rank $(\lambda_i+1)\beta_i$ are the numerical multiples of $X_i^{\lambda_i+1}$. Therefore $z_i = c_i X_i^{\lambda_i+1}$ $(c_i \in C)$. But from Lemma 6 of [5(b)],

$$X_i^{\lambda_i+1}Y_i^{\lambda_i+1} = X_i^{\lambda_i}[X_i, Y_i^{\lambda_i+1}] + X_i^{\lambda_i}Y_i^{\lambda_i+1}X_i$$
$$= \lambda_i X_i^{\lambda_i}Y_i^{\lambda_i}(H_i - \lambda_i) + X_i^{\lambda_i}Y_i^{\lambda_i+1}X_i \equiv 0 \mod \mathfrak{P}_{\Lambda_0} \quad (1 \leqq i \leqq q).$$

Therefore $1 \equiv 0 \mod \mathfrak{P}_{\Lambda_0}$ and so $\mathfrak{B} = \mathfrak{P}_{\Lambda_0}$. But in view of the following lemma this is impossible and therefore $\mathfrak{M}_0 \neq \mathfrak{B}$.

LEMMA 16. *Let Λ be a linear function on $\mathfrak{h}$ and let $\mathfrak{P}_\Lambda$ be the left ideal in $\mathfrak{B}$ defined by* $\mathfrak{P}_\Lambda = \sum_{\beta \in P} \mathfrak{B}X_\beta + \sum_{\beta \in P} \mathfrak{B}(H_\beta - \Lambda(H_\beta))$. *Then* $\mathfrak{P}_\Lambda \neq \mathfrak{B}$.

The proof is similar to that of Lemma 5 of [5(b)]. Let us suppose $\mathfrak{P}_\Lambda = \mathfrak{B}$. Then if $\mathfrak{P} = \sum_{\beta \in P} \mathfrak{B}X_\beta$, we can choose elements $u_\beta \in \mathfrak{B}$ such that

$$1 \equiv \sum_{\beta \in P} u_\beta(H_\beta - \Lambda(H_\beta)) \mod \mathfrak{P}.$$

Again since both 1 and $H_\beta - \Lambda(H_\beta)$ are of rank zero we may assume that u_β is of rank zero. Then $u_\beta(H_\beta - \Lambda(H_\beta)) = (H_\beta - \Lambda(H_\beta))u_\beta$ and $u_\beta \in \mathfrak{B} = \mathfrak{B}\mathfrak{H}\mathfrak{U}$. Since u_β is of rank zero and we are interested only in the congruence mod $\mathfrak{P}$, we may assume that $u_\beta \in \mathfrak{H}$. Hence

$$1 - \sum_{\beta \in P} u_\beta(H_\beta - \Lambda(H_\beta)) \in \mathfrak{P} \cap \mathfrak{H} = \{0\}$$

from Lemma 33 of [5(b)]. Now $\mathfrak{H}$ is the universal enveloping algebra of the abelian Lie algebra $\mathfrak{h}$ (Lemma 21 of [5(b)]). Therefore if ρ is the representation of $\mathfrak{h}$ of degree 1 given by $\rho(H) = \Lambda(H)$ $(H \in \mathfrak{h})$, ρ can be extended uniquely to the representation of $\mathfrak{H}$ such that $\rho(1) = 1$. But since $1 = \sum_{\beta \in P} u_\beta(H_\beta - \Lambda(H_\beta))$ it follows that

$$1 = \rho(1) = \sum_{\beta \in P} \rho(u_\beta)\{\rho(H_\beta) - \Lambda(H_\beta)\} = 0$$

and so we get a contradiction. This proves the lemma.

Since $\mathfrak{M}_0 \neq \mathfrak{B}$, by Zorn's lemma[7] there exists a maximal left ideal $\mathfrak{M}$ in $\mathfrak{B}$ containing $\mathfrak{M}_0$. Let $b \to b^*$ denote the natural mapping and π the natural representation of $\mathfrak{B}$ on $\mathfrak{B}^* = \mathfrak{B}/\mathfrak{M}$. Since $\mathfrak{M}$ is maximal, π is irreducible. Moreover since $\mathfrak{M} \neq \mathfrak{B}$, $1^* \neq 0$ and $\pi(\mathfrak{M})1^* = \{0\}$. Hence if we put $V = \mathfrak{B}^*$ and $\psi = 1^*$ the first two conditions of Theorem 2 are fulfilled. So we have only to prove that $\dim \pi(\mathfrak{X})1^*$ is finite.

For any root β let s_β denote the Weyl reflexion corresponding to β and let W' be the subgroup of the Weyl group W generated by $s_i = s_{\beta_i}$, $1 \leqq i \leqq q$. Since X_i and $Y_i^{\lambda_i+1}$ $(1 \leqq i \leqq q)$ are in $\mathfrak{M}$ it follows easily (by the argument used in the proof of Lemma 6) that for any $b^* \in \mathfrak{B}^*$ we can find an integer

[7] Since $\mathfrak{B}$ satisfies the ascending chain conditions for left ideal (see [2]) it is not necessary to invoke Zorn's lemma.

$r \geqq 0$ such that $\pi(X_i^{r+1})b^* = \pi(Y_i^{r+1})b^* = 0$, $i = 1, \cdots, q$. Hence we can conclude from Lemma 1 that if Λ is a weight of π the same holds for $s\Lambda$ ($s \in W'$). Now let α be a positive compact root. Then α is not totally positive and therefore from Lemma 14, $\alpha = s\beta_i$ for some $s \in W'$ and some i ($1 \leqq i \leqq q$). Therefore $s_\alpha = ss_i s^{-1} \in W'$. Now suppose d is a positive integer such that $\pi(X_{-\alpha}^{d})1^* \neq 0$. Then $\Lambda_0 - d\alpha$ is a weight of π and hence $s_\alpha(\Lambda_0 - d\alpha) = s_\alpha\Lambda_0 + d\alpha$ is also a weight of π. But then it follows from Lemma 2 that $\Lambda_0 - (s_\alpha\Lambda_0 + d\alpha) = (\Lambda_0(H_\alpha) - d)\alpha \geqq 0$ in $\mathfrak{F}_R$ and therefore $d \leqq \Lambda_0(H_\alpha)$. This shows that if $r > \Lambda_0(H_\alpha)$, $\pi(X_{-\alpha}^{r})1^* = 0$. Hence if $\alpha_1, \cdots, \alpha_p$ are all the distinct positive compact roots of $\mathfrak{g}$, we can choose a positive integer m such that $\pi(X_{-\alpha_i}^{m})1^* = 0$ ($1 \leqq i \leqq p$). From this we can conclude (see the proof of Lemma 2) that $\pi(\mathfrak{X})1^*$ is spanned by $\pi(X_{-\alpha_1}^{m_1} \cdots X_{-\alpha_p}^{m_p})1^*$, $0 \leqq m_1, \cdots, m_p \leqq m$. This proves that the dimension of $\pi(\mathfrak{X})1^*$ is finite. Since the last statement of Theorem 2 is an immediate consequence of the corollary to Lemma 5, the proof of this theorem is now complete.

We denote the representation π of Theorem 2 by π_{Λ_0}. Then π_{Λ_0} is uniquely determined by Λ_0 up to equivalence.

7. Infinitesimally unitary representations. For any given linear function Λ on $\mathfrak{h}$, define the left ideal $\mathfrak{P}_\Lambda$ of $\mathfrak{B}$ as in Lemma 16. Then $\mathfrak{P}_\Lambda \neq \mathfrak{B}$ and therefore by Zorn's lemma[7] there exists a maximal left ideal in $\mathfrak{B}$ containing $\mathfrak{P}_\Lambda$. Moreover we know from Lemma 5 that this maximal left ideal is unique. We shall denote it by $\mathfrak{M}_\Lambda$ and the natural representation of $\mathfrak{B}$ on $\mathfrak{B}/\mathfrak{M}_\Lambda$ by π_Λ. In case Λ satisfies the condition of Theorem 2, this definition of π_Λ agrees with the one given in Section 6. We shall say that π_Λ is *infinitesimally unitary* (see [5(c), p. 233]) if it is possible to define a scalar product (a^*, b^*) in $\mathfrak{B}^* = \mathfrak{B}/\mathfrak{M}_\Lambda$ ($a^*, b^* \in \mathfrak{B}^*$) such that under this product $\mathfrak{B}^*$ fulfills all the requirements of a Hilbert space except completeness and $(a^*, \pi_\Lambda(X)b^*) = -(\pi_\Lambda(X)a^*, b^*)$ for all $X \in \mathfrak{g}_0$ and $a^*, b^* \in \mathfrak{B}^*$. We shall now propose to investigate the conditions on Λ under which π_Λ is infinitesimally unitary.

Let $H, \cdots, H_l$ be a base for $\mathfrak{h}_0$ over R and let $x_1, \cdots, x_l$ be l independent (commutative) indeterminates. Let $\mathfrak{H}$ be the subalgebra of $\mathfrak{B}$ generated by $(1, \mathfrak{h})$. Since $\mathfrak{H}$ is the universal enveloping algebra of $\mathfrak{h}$, there exists a uniquely determined homomorphism ξ_x of $\mathfrak{H}$ into $C[x] = C[x_1, \cdots, x_l]$ such that $\xi_x(H_i) = x_i$, $i = 1, \cdots, l$ and $\xi_x(1) = 1$. We shall now extend ξ_x to a linear mapping of $\mathfrak{B}$ into $C[x]$ as follows. Put $\mathfrak{P} = \sum_{\beta \in P} \mathfrak{B}X_\beta$ as before and let $\mathfrak{U}$ and $\mathfrak{V}$ respectively be the subalgebras of $\mathfrak{B}$ generated by $(1, X_\beta)_{\beta \in P}$ and $(1, X_{-\beta})_{\beta \in P}$. Then as we have seen above $\mathfrak{B} = \mathfrak{B}\mathfrak{H}\mathfrak{U}$ and therefore it is

obvious that $\mathfrak{B} = \mathfrak{P} + \mathfrak{B}\mathfrak{H}$. Now every element of $\mathfrak{B}\mathfrak{H}$, which is of rank zero, must lie in $\mathfrak{H}$ and so for every $b \in \mathfrak{B}$ of rank zero we can choose an element $h \in \mathfrak{H}$ such that $b \equiv h \bmod \mathfrak{P}$. Moreover h is unique since $\mathfrak{P} \cap \mathfrak{H} = \{0\}$ from Lemma 33 of [5(b)]. We define $\xi_x(b) = \xi_x(h)$. Now let z be any element in $\mathfrak{B}$ and z_0 its component of rank zero (see Section 2). We extend ξ_x on $\mathfrak{B}$ by setting $\xi_x(z) = \xi_x(z_0)$. If $\mathfrak{Z}$ is the center of $\mathfrak{B}$ and χ_x is the mapping of $\mathfrak{Z}$ into $C[x]$ as defined in [5(b), Lemma 36], it is obvious that $\xi_x(z) = \chi_x(z)$ $(z \in \mathfrak{Z})$.

LEMMA 17. *Let z_1, z_2 be two elements of $\mathfrak{B}$ and suppose at least one of them has rank zero. Then $\xi_x(z_1 z_2) = \xi_x(z_2 z_1) = \xi_x(z_1)\xi_x(z_2)$. Moreover $\xi_x(z) = \chi_x(z)$ if $z \in \mathfrak{Z}$.*

Suppose z_1 is of rank zero and $(z_2)_0$ is the component of z_2 of rank zero. Then $z_1(z_2)_0$ and $(z_2)_0 z_1$ are the components of rank zero of $z_1 z_2$ and $z_2 z_1$ respectively. Let $z_1 = h_1 + p_1$, $(z_2)_0 = h_2 + p_2$ where $h_1, h_2 \in \mathfrak{H}$ and $p_1, p_2 \in \mathfrak{P}$. Then $(z_2)_0 z_1 \equiv z_1(z_2)_0 \equiv h_1 h_2 \bmod \mathfrak{P}$ and this proves our assertion.

Λ being any linear function on $\mathfrak{h}$, we denote by $\xi_\Lambda(b)$ $(b \in \mathfrak{B})$ the value of the polynomial $\xi_x(b)$ under the substitution $x_i = \Lambda(H_i)$, $1 \leqq i \leqq l$. Let η be the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$. We can extend $-\eta$ uniquely to an anti-automorphism ϕ of $\mathfrak{B}$ over R. Then $\phi(ab) = \phi(ba)$ $(a, b \in \mathfrak{B})$ and $\phi(X) = -\eta(X)$ $(X \in \mathfrak{g})$.

LEMMA 18. *Let Λ be a linear function on $\mathfrak{h}$. Then π_Λ is infinitesimally unitary if and only if $\xi_\Lambda(\phi(z)z)$ is real and nonnegative for every $z \in \mathfrak{B}$.*

Suppose $\pi = \pi_\Lambda$ is infinitesimally unitary. Let $a \to a^*$ denote the natural mapping of $\mathfrak{B}$ on $\mathfrak{B}^* = \mathfrak{B}/\mathfrak{M}_\Lambda$ and (a^*, b^*) the corresponding scalar product on $\mathfrak{B}^*$. We assume, as we may, that $(1^*, 1^*) = 1$. Then if $H \in \mathfrak{h}_0$, $(1^*, \pi(H)1^*) = -(\pi(H)1^*, 1^*)$ and therefore $\Lambda(H)$ is purely imaginary. Now suppose z is an element of $\mathfrak{B}$ of rank λ so that $[H, z] = \lambda(H)z$ $(H \in \mathfrak{h})$. Then if $H \in \mathfrak{h}_0$,

$$\begin{aligned}\lambda(H)(1^*, \pi(z)1^*) &= (1^*, \pi([H, z])1^*) \\ &= -(\pi(H)1^*, \pi(z)1^*) - (1^*, \pi(zH)1^*) \\ &= \Lambda(H)(1^*, \pi(z)1^*) - \Lambda(H)(1^*, \pi(z)1^*) = 0\end{aligned}$$

and so if $\lambda \neq 0$, $(1^*, \pi(z)1^*) = 0$. Therefore if b is any element in $\mathfrak{B}$ and b_0 is its component of rank zero, $(1^*, \pi(b)1^*) = (1^*, \pi(b_0)1^*)$. Moreover if $b_0 \equiv h \bmod \mathfrak{P}$ $(h \in \mathfrak{H})$, it is obvious that

$$(1^*, \pi(b)1^*) = (1^*, \pi(h)1^*) = \xi_\Lambda(h) = \xi_\Lambda(b)$$

since $\pi(\mathfrak{P})1^* = \{0\}$. This proves that $(1^*, \pi(b)1^*) = \xi_\Lambda(b)$ for all $b \in \mathfrak{B}$. On the other hand $(\pi(X)a^*, b^*) = (a^*, \pi(\phi(X))b^*)$ $(a^*, b^* \in \mathfrak{B}, X \in \mathfrak{g})$ and so it is clear that $(\pi(z)a^*, b^*) = (a^*, \pi(\phi(z)b^*)$ $(z \in \mathfrak{B})$. Hence

$$\xi_\Lambda(\phi(z)z) = (1^*, \pi(\phi(z)z)1^*) = (\pi(z)1^*, \pi(z)1^*) \geqq 0.$$

This shows that our condition is necessary.

Conversely suppose $\xi_\Lambda(\phi(z)z)$ is real and nonnegative for all $z \in \mathfrak{B}$. Put $(w, z) = \xi_\Lambda(\phi(w)z)$ $(w, z \in \mathfrak{B})$ and $|z| = (z, z)^{\frac{1}{2}} \geqq 0$. Then a well-known elementary argument shows that $|(w, z)| \leqq |w||z|$ and $|z + w| \leqq |z| + |w|$. Let $\mathfrak{M}$ be the set of all elements $z \in \mathfrak{B}$ such that $|z| = 0$. Since $|z_1 + z_2| \leqq |z_1| + |z_2|$, $\mathfrak{M}$ is a vector space. Moreover if $b, z, w \in \mathfrak{B}$ and $b' = \phi(b)$,

$$(w, bz) = \xi_\Lambda(\phi(w)bz) = \xi_\Lambda(\phi(b'w)z) = (b'w, z).$$

Hence if $|z| = 0$, $|(w, bz)| \leqq |b'w||z| = 0$ and therefore in particular $|bz|^2 = 0$. This proves that $\mathfrak{M}$ is a left ideal. Let $b \to b^*$ denote the natural mapping and π the natural representation of $\mathfrak{B}$ on $\mathfrak{B}^* = \mathfrak{B}/\mathfrak{M}$. It is clear that (w, z) depends only on w^* and z^*. Put $(w^*, z^*) = (w, z)$ and $|z^*| = |z|$. Then $|z^*| \neq 0$ unless $z^* = 0$ and if $X \in \mathfrak{g}_0$,

$$\begin{aligned}(w^*, \pi(X)z^*) = (w, Xz) = \xi_\Lambda(\phi(w)Xz) = -\xi_\Lambda(\phi(Xw)z)\\ = -(\pi(X)w^*, z^*) \qquad (w, z \in \mathfrak{B}).\end{aligned}$$

Hence in order to prove the lemma it only remains to prove that $\mathfrak{M} = \mathfrak{M}_\Lambda$. It is clear that $\xi_\Lambda(z) = 0$ if $z \in \mathfrak{P}$ and therefore it follows from Lemma 17 that $\xi_\Lambda(\mathfrak{P}_\Lambda) = \{0\}$. But since $\mathfrak{P}_\Lambda$ is a left ideal this implies that $\mathfrak{M} \supset \mathfrak{P}_\Lambda$. Hence it would be enough to prove that $\mathfrak{M}$ is maximal. Since $\xi_\Lambda(1) = 1 \neq 0$, $\mathfrak{M} \neq \mathfrak{B}$. Now suppose $\mathfrak{M}'$ is a left ideal in $\mathfrak{B}$ such that $\mathfrak{B} \neq \mathfrak{M}'$ and $\mathfrak{M}' \supset \mathfrak{M}$. Then if $b' \in \mathfrak{M}'$ and $H \in \mathfrak{h}$

$$[H, b'] = Hb' - b'(H - \Lambda(H)) - \Lambda(H)b' \in \mathfrak{M}'$$

and from this we can conclude that all the components of b' of various ranks also lie in $\mathfrak{M}'$. Let $(b')_0$ be its component of rank zero and choose $h \in \mathfrak{H}$ such that $(b')_0 \equiv h \bmod \mathfrak{P}$. Then it is obvious that $(b')_0 \equiv h \equiv \xi_\Lambda(h) \bmod \mathfrak{P}_\Lambda$. Since $(b')_0$ and $\mathfrak{P}_\Lambda$ are both contained in $\mathfrak{M}'$, $\xi_\Lambda(b') = \xi_\Lambda(h) \in \mathfrak{M}'$. However $\mathfrak{M}' \neq \mathfrak{B}$ and so this implies that $\xi_\Lambda(b') = 0$. Now put $b' = \phi(z')z'$ where z' is any element in $\mathfrak{M}'$. Then $b' \in \mathfrak{M}'$ and therefore $(z', z') = \xi_\Lambda(\phi(z')z') = 0$. This shows that $z' \in \mathfrak{M}$ and hence $\mathfrak{M}' = \mathfrak{M}$. Our proof is now complete.

Notice that it follows automatically from the above condition on ξ_Λ that $\dim \pi(\mathfrak{X})1^*$ is finite. For this it would be enough to verify (see Lemma 2)

that if α is a positive compact root of $\mathfrak{g}$, $\pi(X_{-\alpha}{}^r)1^* = 0$ for some integer $r > 0$. But $\eta(X_\alpha) = -X_{-\alpha}$ (see Section 4) and therefore

$$\xi_\Lambda(\phi(X_{-\alpha}{}^r)X_{-\alpha}{}^r) = \xi_\Lambda(X_\alpha{}^r X_{-\alpha}{}^r) \geqq 0 \qquad (r \geqq 1).$$

Moreover we know from Lemma 6 of [5(b)] that

$$\begin{aligned} X_\alpha{}^r X_\alpha{}^{-r} &= X_\alpha{}^{r-1}[X_\alpha, X_{-\alpha}{}^r] + X_\alpha{}^{r-1}X_{-\alpha}{}^r X_\alpha \\ &\equiv r X_\alpha{}^{r-1}X_{-\alpha}{}^{r-1}(\Lambda(H_\alpha) - r + 1) \bmod \mathfrak{P}_\Lambda \qquad (r \geqq 1). \end{aligned}$$

Therefore if $\lambda_\alpha = \Lambda(H_\alpha)$,

$$0 \leqq \xi_\Lambda(X_\alpha{}^r X_{-\alpha}{}^r) = r! \prod_{1 \leqq j \leqq r} (\lambda_\alpha - j + 1)$$

by induction on r. Putting $r = 1$ we conclude that $\lambda_\alpha \geqq 0$. Now let r be the least integer such that $\lambda_\alpha - r + 1 \leqq 0$. Then the above product is negative unless $\lambda_\alpha = r - 1$. Hence we conclude that λ_α is a nonnegative integer and therefore $\xi_\Lambda(X_\alpha{}^r X_{-\alpha}{}^r) = 0$ for $r \geqq \lambda_\alpha + 1$. But since

$$\xi_\Lambda(X_\alpha{}^r X_{-\alpha}{}^r) = |\pi(X_{-\alpha}{}^r)1^*|^2,$$

$\pi(X_{-\alpha}{}^r)1^* = 0$ for $r \geqq \lambda_\alpha + 1$. This shows that π then satisfies all the requirements of Theorem 1 and therefore $\Lambda(H_\beta)$ must be a nonnegative integer for all positive roots β which are not totally positive.

Now suppose β is a noncompact root. Then $\eta(X_{-\beta}) = X_\beta$ (see Section 4) and therefore

$$\xi_\Lambda(\phi(X_{-\beta}{}^r)X_{-\beta}{}^r) = (-1)^r \xi_\Lambda(X_\beta{}^r X_{-\beta}{}^r) = r! \prod_{1 \leqq j \leqq r} (j - 1 - \lambda_\beta) \quad (r \geqq 1)$$

where $\lambda_\beta = \Lambda(H_\beta)$. Since again this product must always be real and nonnegative, we conclude that λ_β is real and $\leqq 0$. In particular if β is not totally positive it follows, in view of our earlier remarks, that $\lambda_\beta = 0$. Thus we have obtained the following corollary.

COROLLARY 1. *In order that π_Λ be infinitesimally unitary it is necessary that Λ should satisfy the following conditions:*

(1) *$\Lambda(H_\alpha)$ is a nonnegative integer for every positive compact root α.*

(2) *$\Lambda(H_\beta) = 0$ for every noncompact positive root β which is not totally positive.*

(3) *$\Lambda(H_\gamma)$ is real and $\leqq 0$ for every totally positive root γ.*

The space $\mathfrak{F}$ of linear functions on $\mathfrak{h}$, being a finite-dimensional vector space over C, has a natural topology.

COROLLARY 2. *The set of all linear functions* Λ *on* $\mathfrak{h}$ *for which* π_Λ *is infinitesimally unitary, is closed in* $\mathfrak{F}$.

For a fixed element $b \in \mathfrak{B}$, $\Lambda \to \xi_\Lambda(b)$ $(\Lambda \in \mathfrak{F})$ is obviously a polynomial function on $\mathfrak{F}$. Hence it is continuous and therefore the corollary follows from Lemma 18.

Define $\mathfrak{g}'$ as in the proof of Lemma 13.

LEMMA 19. *Suppose* Λ *satisfies the conditions of Theorem 2 and* $\Lambda(H_\beta) = 0$ *for every positive root* β *which is not totally positive. Then* $\pi_\Lambda(X) = 0$ *for all* $X \in \mathfrak{g}'$.

We use the notation of the proof of Lemma 13. Let α be a positive compact root of $\mathfrak{g}'$. Then we can choose a noncompact root β of $\mathfrak{g}'$ such that $\beta + \alpha$ is a root. For otherwise $[X_\alpha, X_\beta] = 0$ for every noncompact root β of $\mathfrak{g}'$. Since it is obvious that $\mathfrak{q} = \mathfrak{g}' \cap \mathfrak{p}$ is spanned by the elements X_β corresponding to all noncompact roots β of $\mathfrak{g}'$, X_α lies in the centralizer $\mathfrak{l}$ of $\mathfrak{q}$ in $\mathfrak{k}$. But $\mathfrak{l} \subset \mathfrak{g}_+$ and therefore $X_\alpha \in \mathfrak{g}_+ \cap \mathfrak{g}' = \{0\}$, which of course is false. So let β be a noncompact root of $\mathfrak{g}'$ such that $\gamma = \beta + \alpha$ is a root. Since $\mathfrak{q} \cap (\mathfrak{p}_+ + \mathfrak{p}_-) = \{0\}$, no root of $\mathfrak{g}'$ can be totally positive or totally negative. Therefore $\Lambda(H_\beta) = \Lambda(H_\gamma) = 0$. But since $\alpha = \gamma - \beta$, H_α is a linear combination of H_β and H_γ and so $\Lambda(H_\alpha) = 0$. This shows that $\Lambda(H_\beta) = 0$ for every root β of $\mathfrak{g}'$. Moreover if β is a positive root of $\mathfrak{g}'$, it follows from Theorem 1 that $\pi_\Lambda(X_{-\beta})1^* = 0$ since β is not totally positive. Therefore $\pi_\Lambda(\mathfrak{g}')1^* = \{0\}$ and since $\mathfrak{g}'$ is an ideal in $\mathfrak{g}$ and π_Λ is irreducible, $\pi_\Lambda(\mathfrak{g}') = \{0\}$.

COROLLARY 1. *If* π_Λ *is infinitesimally unitary* $\pi_\Lambda(\mathfrak{g}') = \{0\}$.

This is an immediate consequence of the above lemma and Corollary 1 to Lemma 18.

COROLLARY 2. *Suppose* $\mathfrak{g}_0$ *is simple and* $\mathfrak{g}_0 \neq \mathfrak{k}_0$. *Then if* $\mathfrak{c}_0 = \{0\}$, π_Λ *cannot be infinitesimally unitary unless* $\Lambda = 0$ *in which case it is the zero representation of* $\mathfrak{g}$ *of degree* 1.

For then we know that $\mathfrak{g} = \mathfrak{g}'$ (see the proof of Corollary 2 to Lemma 13). Now apply Corollary 1 above.

In Corollary 1 to Lemma 18 we have obtained certain necessary conditions on Λ in order that π_Λ may be infinitesimally unitary. It is natural to ask how far these conditions are also sufficient. The following result gives a partial answer to this question.

THEOREM 3. *Let 2ρ denote the sum of all positive roots of $\mathfrak{g}$ and let Λ be a linear function on $\mathfrak{h}$ satisfying the following conditions:*

(1) $\Lambda(H_\alpha)$ *is a nonnegative integer for every positive compact root α.*

(2) $\Lambda(H_\beta) = 0$ *for every noncompact positive root β which is not totally positive.*

(3) $\Lambda(H_\gamma) + \rho(H_\gamma)$ *is real and $\leqq 0$ for every totally positive root γ. Then π_Λ is infinitesimally unitary.*

We leave the proof of this theorem to a subsequent paper (see however [5(h)]).

8. Representations of the group. Let G be the simply connected Lie group with the Lie algebra $\mathfrak{g}_0$ and let K be its analytic subgroup corresponding to $\mathfrak{k}_0$. Let Λ be a linear function on $\mathfrak{h}$ which fulfills the conditions of Theorem 2. We shall now show that π_Λ can be 'extended' to a representation of G on a Hilbert space. Let Ω be the set of all equivalence classes of finite-dimensional irreducible representations of K. Henceforth we shall make free use of the terminology of [5(c)].

THEOREM 4. *Let Λ be a linear function on $\mathfrak{h}$ such that $\Lambda(H_\beta)$ is a nonnegative integer for every positive root β which is not totally positive. Then there exists an irreducible quasi-simple representation σ of G on a Hilbert space $\mathfrak{H}$ with the following property. Let $\mathfrak{H}_\mathfrak{D}$ $(\mathfrak{D} \in \Omega)$ denote the subspace of $\mathfrak{H}$ consisting of all those elements which transform under $\sigma(K)$ according to $\mathfrak{D}$. Then the corresopnding representation σ_0 of $\mathfrak{B}$ defined on $\mathfrak{H}_0 = \sum_{\mathfrak{D} \in \Omega} \mathfrak{H}_\mathfrak{D}$ is equivalent to π_Λ.*

Put $\pi = \pi_\Lambda$ and define $\mathfrak{P}_\Lambda$ and $\mathfrak{M}_\Lambda$ as in Section 7. If z is any element in $\mathfrak{B}$ of rank zero, $z \equiv \xi_\Lambda(z) \bmod \mathfrak{P}_\Lambda$ and therefore in particular if $z \in \mathfrak{z}$, $z - \xi_\Lambda(z) \in \mathfrak{M}_\Lambda$. Moreover we know from Theorem 2 that $\dim \mathfrak{X}/\mathfrak{X} \cap \mathfrak{M}_\Lambda$ is finite and so we can conclude from Lemma 9 of [5(c)] that π is a quasi-simple[8] irreducible representation of $\mathfrak{B}$.

Let $(\beta_1, \cdots, \beta_l)$ and $(\beta_1, \cdots, \beta_q)$ be defined as in the proof of Theorem 2 and let Λ' be the linear function on $\mathfrak{h}$ given by $\Lambda'(H_{\beta_i}) = \Lambda(H_{\beta_i})$, $1 \leqq i \leqq q$ and $\Lambda'(H_{\beta_i}) = 0$, $q < i \leqq l$. Then from Theorem 1 of [5(b)], there exists a finite-dimensional irreducible representation ρ of $\mathfrak{g}$ with the highest weight Λ'. Let U be the representation space of ρ and ϕ a nonzero vector in U belonging to the weight Λ'. Then if α is any positive compact

[8] See [5(d), p. 56] for the definition of quasi-simplicity in this case.

root of $\mathfrak{g}$, $\rho(X_\alpha)\phi = 0$ and $\rho(H_\alpha)\phi = \Lambda'(H_\alpha)\phi = \Lambda(H_\alpha)\phi$ since α is a linear combination of $\beta_1, \cdots, \beta_q$ (Lemma 13). Similarly if $\psi \neq 0$ is a vector belonging to the weight Λ of π, $\pi(X_\alpha)\psi = 0$ and $\pi(H_\alpha)\psi = \Lambda(H_\alpha)\psi$. Let $\mathfrak{X}'$ be the subalgebra of $\mathfrak{B}$ generated by $(1\,\mathfrak{k}')$. Since $\mathfrak{k}'$ is semisimple we can apply Lemma 2 to $\mathfrak{X}'$ and conclude that the representations of $\mathfrak{k}'$ defined on $\pi(\mathfrak{X}')\psi$ and $\rho(\mathfrak{X}')\phi$ are both irreducible and therefore from Lemma 5 they are equivalent. Since ρ is a finite-dimensional representation of $\mathfrak{g}$, the existence of σ now follows from Theorem 4 of [5(d)].

COROLLARY. *Let λ, μ be two linear functions on $\mathfrak{h}$. Then if π_λ and π_μ are both infinitesimally unitary the same holds for $\pi_{\lambda+\mu}$.*

Suppose π_λ and π_μ are infinitesimally unitary. Then λ and μ both satisfy the conditions of Theorem 2 (see Corollary 1 to Lemma 18) and therefore from Theorem 4 above and Theorem 9 of [5(c)] we can find two irreducible unitary representations σ_λ, σ_μ of G on the Hilbert spaces $\mathfrak{H}_\lambda$, $\mathfrak{H}_\mu$ respectively such that the corresponding representations of $\mathfrak{B}$ defined on $\mathfrak{H}_\lambda^0 = \sum_{\mathfrak{D}\in\Omega} (\mathfrak{H}_\lambda)_\mathfrak{D}$ and $\mathfrak{H}_\mu^0 = \sum_{\mathfrak{D}\in\Omega} (\mathfrak{H}_\mu)_\mathfrak{D}$ are equivalent to π_λ and π_μ. We denote these representations of $\mathfrak{B}$ by σ_λ^0 and σ_μ^0. Then it is clear from the proofs of Lemma 18 and Theorem 9 of [5(c)] that we can choose $\psi_\lambda \in \mathfrak{H}_\lambda^0$ and $\psi_\mu \in \mathfrak{H}_\mu^0$ such that

$$(\psi_\lambda, \sigma_\lambda^0(z)\psi_\lambda) = \xi_\lambda(z), \qquad (\psi_\mu, \sigma_\mu^0(z)\psi_\mu) = \xi_\mu(z)$$

for all $z \in \mathfrak{B}$. Let $\mathfrak{H}$ be the Hilbert space which is the Kronecker product of $\mathfrak{H}_\lambda$ and $\mathfrak{H}_\mu$ (in the sense of Hilbert space theory) and let $\psi = \psi_\lambda \times \psi_\mu$. Then if ψ' and ψ'' are well-behaved under σ_λ and σ_μ respectively it is obvious that $\psi' \times \psi''$ is well-behaved under the representation $\sigma = \sigma_\lambda \times \sigma_\mu$ of G on $\mathfrak{H}$. Let σ^0 denote the representation of $\mathfrak{B}$ defined under σ on the subspace $\mathfrak{H}^0$ consisting of all well-behaved elements in $\mathfrak{H}$. Then it is obvious that $\sigma^0(X)(\phi_1 \times \phi_2) = (\sigma_\lambda^0(X)\phi_1) \times \phi_2 + \phi_1 \times (\sigma_\mu^0(X)\phi_2)$ $(X \in \mathfrak{g}, \phi_1 \in \mathfrak{H}_\lambda^0, \phi_2 \in \mathfrak{H}_\mu^0)$ and therefore $\sigma^0(X_\beta)\psi = 0$ for all positive roots β and $\sigma^0(H)\psi = (\lambda(H) + \mu(H))\psi$ if $H \in \mathfrak{h}$. Hence $\sigma^0(\mathfrak{P}_{\lambda+\mu})\psi = \{0\}$ in the notation of Section 7. On the other hand since σ is a unitary representation of G,

$$(\phi', \sigma^0(X)\phi) = -(\sigma^0(X)\phi', \phi) \qquad (X \in \mathfrak{g}_0; \phi, \phi' \in \mathfrak{H}^0).$$

Therefore if z is any element in $\mathfrak{B}$ of rank $\nu \neq 0$ it follows (see the proof of Lemma 18) that $(\psi, \sigma^0(z)\psi) = 0$. This shows that

$$(\psi, \sigma^0(b)\psi) = \xi_{\lambda+\mu}(b)\,|\,\psi\,|^2 \qquad (b \in \mathfrak{B})$$

and therefore

$$| \sigma^0(z)\psi |^2 = (\sigma^0(z)\psi, \sigma^0(z)\psi) = (\psi, \sigma^0(\phi(z)z)\psi) = \xi_{\lambda+\mu}(\phi(z)z) | \psi |^2$$

for any $z \in \mathfrak{B}$. Hence $\pi_{\lambda+\mu}$ is infinitesimally unitary from Lemma 18.

9. An auxiliary result. In this section we shall prove a result which will be required in another paper. Let Λ_0 be a linear function on $\mathfrak{h}$ satisfying the conditions of Theorem 3. As before let P be the set of all positive roots and P_k and P_+ respectively the subsets of P consisting of all compact and all totally positive roots. Let P' be the complement of $P_k \cup P_+$ in P. We know (Lemma 5) that there exists a unique left ideal $\mathfrak{Y}_{\Lambda_0}$ in $\mathfrak{X}$ containing

$$\sum_{\alpha \in P_k} \mathfrak{X} X_\alpha + \sum_{\gamma \in P} \mathfrak{X}(H_\gamma - \Lambda_0(H_\gamma)).$$

Our object now is to prove the following theorem.

THEOREM 5. *Let Λ_0 be a linear function on $\mathfrak{h}$ satisfying the conditions of Theorem 3. Then*

$$\mathfrak{M} = \sum_{\gamma \in P} \mathfrak{B} X_\gamma + \sum_{\gamma \in P'} \mathfrak{B} X_{-\gamma} + \mathfrak{B}\mathfrak{Y}_{\Lambda_0}$$

is a maximal left ideal in $\mathfrak{B}$.

It is obvious that Λ_0 fulfills the conditions of Theorem 2 and therefore we can consider the representation π_{Λ_0} and choose a vector $\psi \neq 0$ in the representation space corresponding to Theorem 2. Define $\mathfrak{g}_+$ and $\mathfrak{g}'$ as in Section 5. Then we know from Lemma 19 that $\pi_{\Lambda_0}(\mathfrak{g}') = \{0\}$. Moreover it follows from Lemmas 2 and 5 (applied to $\mathfrak{k}$) that $\pi_{\Lambda_0}(\mathfrak{Y}_{\Lambda_0})\psi = \{0\}$. This shows that $\pi_{\Lambda_0}(\mathfrak{M})\psi = \{0\}$ and therefore $\mathfrak{M} \neq \mathfrak{B}$ since $\psi \neq 0$.

Let $\mathfrak{M}_1$ be the set of all elements $b \in \mathfrak{B}$ such that $\pi_{\Lambda_0}(b)\psi = 0$. Since $\mathfrak{Y}_{\Lambda_0}$ is a maximal left ideal in $\mathfrak{X}$, it is clear that $\mathfrak{M}_1 \cap \mathfrak{X} = \mathfrak{Y}_{\Lambda_0}$. Therefore in view of the fact that $\mathfrak{g}' \subset \mathfrak{M}_1$, we can conclude that $\mathfrak{g}' \cap \mathfrak{k} \subset \mathfrak{Y}_{\Lambda_0}$. On the other hand $\mathfrak{g}' = \sum_{\gamma \in P'} (CX_\gamma + CX_{-\gamma}) + \mathfrak{g}' \cap \mathfrak{k}$ and so it follows that $\mathfrak{g}' \subset \mathfrak{M}$.

Let $b \to b^*$ $(b \in \mathfrak{B})$ denote the natural mapping and π the natural representation of $\mathfrak{B}$ on $\mathfrak{B}^* = \mathfrak{B}/\mathfrak{M}$. In order to prove that $\mathfrak{M}$ is a maximal left ideal in $\mathfrak{B}$ it is sufficient to show that π is irreducible. Let $\mathfrak{B}_+$ be the subalgebra of $\mathfrak{B}$ generated by $(1, \mathfrak{g}_+)$. Since $\mathfrak{g}'$ is an ideal in $\mathfrak{g}$ and $\mathfrak{g}' \subset \mathfrak{M}$, it follows that $\pi(\mathfrak{g}') = \{0\}$ and therefore $\mathfrak{B}^* = \pi(\mathfrak{B}_+)1^*$. Hence it would be enough to prove that $\mathfrak{B}_+ \cap \mathfrak{M}$ is a maximal left ideal in $\mathfrak{B}_+$. Let $\mathfrak{X}_+$ be the subalgebra of $\mathfrak{X}$ generated by $(1, \mathfrak{k} \cap \mathfrak{g}_+)$. Since $\pi(\mathfrak{Y}_{\Lambda_0})1^* = \{0\}$ and since $\mathfrak{Y}_{\Lambda_0}$ is a maximal left ideal in $\mathfrak{X}$, it is clear that $\mathfrak{X}^* = \pi(\mathfrak{X})1^*$ is

irreducible under $\pi(\mathfrak{X})$ and therefore also under $\pi(\mathfrak{X}_+)$. Hence $\mathfrak{Y}_+ = \mathfrak{X}_+ \cap \mathfrak{Y}_{\Lambda_0}$ is a maximal left ideal in $\mathfrak{X}_+$. Let Q denote the set of all positive roots of $\mathfrak{g}_+$ (see Section 5). Then from Lemma 5, $\mathfrak{Y}_+$ must be the unique maximal left ideal in $\mathfrak{X}_+$ containing

$$\sum_{\gamma \in Q_k} \mathfrak{X}_+ X_\alpha + \sum_{\gamma \in Q} \mathfrak{X}_+ (H_\gamma - \Lambda_0(H_\gamma))$$

where $Q_k = Q \cap P_k$. Now if $\mathfrak{M}_+ = \sum_{\gamma \in Q} \mathfrak{B}_+ X_\gamma + \mathfrak{B}_+ \mathfrak{Y}_+$ it is obvious that $\mathfrak{B}_+ \cap \mathfrak{M} \supset \mathfrak{M}_+$ and therefore it is enough to prove that $\mathfrak{M}_+$ is a maximal left ideal in $\mathfrak{B}_+$. Let $2\rho_+ = \sum_{\gamma \in Q} \gamma$ and $2\rho = \sum_{\gamma \in P} \gamma$. If β is a root in $\mathfrak{c}'$ and γ a root in $\mathfrak{g}_+$ then $\beta(H_\gamma) = 0$. Hence it follows that $\rho(H_\gamma) = \rho_+(H_\gamma)$ for all $\gamma \in Q$. This shows (by comparison with the conditions of Theorem 3) that the problem of proving the maximality of $\mathfrak{M}_+$ in $\mathfrak{B}_+$ is exactly the same as our original problem if we replace $\mathfrak{g}$ by $\mathfrak{g}_+$. Therefore we may assume that $\mathfrak{g}' = \{0\}$ and $P = P_k \cup P_+$.

Let $\mathfrak{M}_1 \neq \mathfrak{M}$ be a left ideal in $\mathfrak{B}$ containing $\mathfrak{M}$. We have to show that $\mathfrak{M}_1 = \mathfrak{B}$. For any linear function Λ on $\mathfrak{h}$, let $\mathfrak{B}_\Lambda{}^*$ denote the set of elements $b^* \in \mathfrak{B}^*$ such that $\pi(H)b^* = \Lambda(H)b^*$ $(H \in \mathfrak{h})$. Then it is clear that $\mathfrak{B}^* = \sum_\Lambda \mathfrak{B}_\Lambda{}^*$ and $\mathfrak{B}_\Lambda{}^* = \{0\}$ unless $\Lambda_0 - \Lambda$ is a linear combination of positive roots with coefficients which are all nonnegative integers (see the proof of Lemma 2). Moreover $\mathfrak{M}_1{}^* = \sum_\Lambda (\mathfrak{M}_1{}^* \cap \mathfrak{B}_\Lambda{}^*) \neq \{0\}$ since $\mathfrak{M}_1 \neq \mathfrak{M}$ (Lemma1 of [5(b)]). Hence in view of the above remarks, there exists a highest weight Λ_1 such that $\mathfrak{M}_1{}^* \cap \mathfrak{B}_{\Lambda_1}{}^* \neq \{0\}$. Choose an element $\phi \neq 0$ in $\mathfrak{M}_1{}^* \cap \mathfrak{B}_{\Lambda_1}{}^*$. Then if γ is a positive root $\mathfrak{M}_1{}^* \cap \mathfrak{B}_{\Lambda_1+\gamma}{}^* = \{0\}$ and therefore $\pi(X_\gamma)\phi = 0$. Now define $\chi_x(z)$ $(z \in \mathfrak{Z})$ as in Lemma 17 and for any linear function λ on $\mathfrak{h}$ let $\chi_\lambda(z)$ denote the value of this polynomial under the substitution $x_i = \lambda(H_i)$ $(1 \leqq i \leqq l)$. (Here we are using the notation of Section 7). Then it is clear that $\pi(z)\phi = \chi_{\Lambda_1}(z)\phi$ for all $z \in \mathfrak{Z}$. Similarly $\pi(z)1^* = \chi_{\Lambda_0}(z)1^*$ $(z \in \mathfrak{Z})$ and therefore $\pi(z)b^* = \chi_{\Lambda_0}(z)b^*$ for every $b^* \in \mathfrak{B}^*$. Since $\phi \neq 0$, we conclude that $\chi_{\Lambda_0}(z) = \chi_{\Lambda_1}(z)$ for any $z \in \mathfrak{Z}$. But then if W is the Weyl group of $\mathfrak{g}$ with respect to $\mathfrak{h}$, it follows from Theorem 5 of [5(b)] that $\Lambda_1 + \rho = s_0(\Lambda_0 + \rho)$ for some $s_0 \in W$. (Here $2\rho = \sum_{\gamma \in P} \gamma$).

Let N_- denote the set of all totally negative roots (Section 4) and put $\bar{P} = P_k \cup N_-$. Since $\mathfrak{g}' = \{0\}$, it follows from Lemmas 11 and 12 that $\bar{P}$ satisfies the conditions of Corollary 1 to Lemma 4. Put $\Lambda_0' = \Lambda_0 + \rho$ and let $\mathfrak{F}_R(\bar{P})$ denote the set of those linear functions on $\mathfrak{h}$ which are linear combinations of roots in $\bar{P}$ with real nonnegative coefficients. Since

$\rho(H_\beta) > 0$ for every $\beta \in P$ (see Weyl [7]) it follows from our assumptions on Λ_0 that $\Lambda_0'(H_\gamma) \geqq 0$ for all $\gamma \in P$. As we shall see presently (Lemma 20 below), this implies that $\Lambda_0' - s\Lambda_0' \in \mathfrak{F}^+(\bar{P})$ for every $s \in W$. Let $(\beta_1, \cdots, \beta_l)$ be a fundamental system of roots in P. Suppose $(\beta_1, \cdots, \beta_t)$ are all the totally positive roots and $(\alpha_1, \cdots, \alpha_r)$ all the (distinct) compact roots among $(\beta_1, \cdots, \beta_l)$. Then we know from Lemma 13 that every totally negative root is of the form $-(\beta_i + m_1\alpha_1 + \cdots + m_r\alpha_r)$ where $1 \leqq i \leqq t$ and $m_1, \cdots, m_r \geqq 0$. Hence

$$\begin{aligned}\Lambda_0 - \Lambda_1 &= (\Lambda_0 + \rho) - s_0(\Lambda_0 + \rho) \\ &= -(\mu_1\beta_1 + \cdots + \mu_t\beta_t) + m_1'\alpha_1 + \cdots + m_r'\alpha_r\end{aligned}$$

where $\mu_1, \cdots, \mu_t \geqq 0$. On the other hand $\mathfrak{B}_{\Lambda_1}^* \neq \{0\}$ and therefore, as we have seen above,

$$\Lambda_0 - \Lambda_1 = \nu_1\beta_1 + \cdots + \nu_t\beta_t + n_1\alpha_1 + \cdots + n_r\alpha_r$$

where $\nu_1, \cdots, \nu_t, n_1, \cdots, n_r \geqq 0$. Since $(\beta_1, \cdots, \beta_l)$ are linearly independent we conclude that $-\mu_i = \nu_i = 0$ $(1 \leqq i \leqq t)$ and therefore $\Lambda_0 - \Lambda_1$ is a linear combination of compact roots. Let $\gamma_1, \cdots, \gamma_p$ be all the (distinct) totally positive roots and $\delta_1, \cdots, \delta_q$ all the compact roots in P. Then $\mathfrak{B}^*$ is spanned by elements of the form $\pi(X_{-\gamma_1}^{\mu_1} \cdots X_{-\gamma_p}^{\mu_p} X_{-\delta_1}^{\lambda_1} \cdots X_{-\delta_1}^{\lambda_q})1^*$ where λ_i and μ_i are nonnegative integers (see the proof of Lemma 2). Since $\phi \in \mathfrak{B}_{\Lambda_1}^*$, it must be a linear combination of those elements of the above form for which

$$\mu_1\gamma_1 + \cdots + \mu_p\gamma_p + \lambda_1\delta_1 + \cdots + \lambda_q\delta_q = \Lambda_0 - \Lambda_1.$$

However since $\Lambda_0 - \Lambda_1$ is a linear combination of compact roots it follows from Lemma 13 that $\mu_1 = \cdots = \mu_p = 0$ and therefore $\phi \in \pi(\mathfrak{X})1^* = \mathfrak{X}^*$. But $\pi(X_\alpha)\phi = 0$ $(\alpha \in P_k)$ and as we have seen above $\mathfrak{X}^*$ is irreducible under $\pi(\mathfrak{X})$. Therefore it follows from Lemma 2 (applied to $\mathfrak{X}$ and ϕ) that $\Lambda_1 - \Lambda_0 \geqq 0$. But $\Lambda_0 - \Lambda_1 = \lambda_1\delta_1 + \cdots + \lambda_q\delta_q \geqq 0$. Hence $\Lambda_0 = \Lambda_1$ and $\lambda_1 = \cdots = \lambda_q = 0$. This shows that $\phi = c1^*$ $(c \in C)$ and since $\phi \neq 0$ it follows that $\mathfrak{M}_1^* \supset \pi(\mathfrak{B})\phi = \mathfrak{B}^*$ and therefore $\mathfrak{M}_1 = \mathfrak{B}$. Hence in order to complete the proof it remains to prove the following lemma.

LEMMA 20. *Let $\bar{P}_0$ be a fundamental system of roots of $\mathfrak{g}$ and let $\mathfrak{F}^+(\bar{P}_0)$ denote the set of all linear functions on $\mathfrak{h}$ which can be written as linear combinations of roots in $\bar{P}_0$ with real nonnegative coefficients. Then if $\Lambda \in \mathfrak{F}_R$ and $\Lambda(H_\beta) \geqq 0$ for all $\beta \in \bar{P}_0$, $\Lambda - s\Lambda \in \mathfrak{F}^+(\bar{P}_0)$ for every $s \in W$.*

It is enough to prove this lemma under the additional assumption that

$\Lambda(H_\beta)$ $(\beta \in P_0)$ are all rational since the general case then follows immediately by continuity. Moreover by multiplying Λ by a suitable positive integer, the problem may be reduced to the case when $\Lambda(H_\beta)$ $(\beta \in \bar{P}_0)$ are all nonnegative integers. But in this case we know (Theorem 1 of [5(b)] that there exists a finite-dimensional irreducible representation π of $\mathfrak{g}$ such that Λ is a weight of π while $\Lambda + \gamma$ is not a weight for any root $\gamma \in \mathfrak{F}^+(\bar{P}_0)$. Then it follows from Lemma 2 that if Λ' is any other weight of π, $\Lambda - \Lambda' \in \mathfrak{F}^+(\bar{P}_0)$. Therefore since $s\Lambda$ is a weight for every $s \in W$, we get the result.

Let us keep to the notation of the proof of Theorem 5 and put $\mathfrak{p}_- = \sum_{\gamma \in P_+} CX_{-\gamma}$. Since $[\mathfrak{p}_-, \mathfrak{p}_-] = \{0\}$ (Lemma 11) the subalgebra $\mathfrak{P}_-$ of $\mathfrak{B}$ generated by $(1, \mathfrak{p}_-)$ is abelian. We consider the tensor product $\mathfrak{P}_- \times \mathfrak{X}^*$.

LEMMA 21. *Let Γ denote the linear mapping $\mathfrak{P}_- \times \mathfrak{X}^*$ into $\mathfrak{B}^*$ such that $\Gamma(p \times x^*) = \pi(p)x^*$ $(p \in \mathfrak{P}_-, x^* \in \mathfrak{X}^*)$. Then Γ is a linear isomorphism of $\mathfrak{P}_- \times \mathfrak{X}^*$ onto $\mathfrak{B}^*$.*

We have seen during the proof of Theorem 4 that $\mathfrak{B}^* = \pi(\mathfrak{B}_+)1^*$ and $\mathfrak{X}^* = \pi(\mathfrak{X}_+)1^*$. Also we proved that $\mathfrak{M}_+ = \sum_{\gamma \in Q} \mathfrak{B}X_\gamma + \mathfrak{B}_+ Y_+$ is a maximal left ideal in $\mathfrak{B}_+$. Therefore $\mathfrak{B} = \mathfrak{B}_+ + \mathfrak{M}$, $\mathfrak{X} = \mathfrak{X}_+ + \mathfrak{Y}_{\Lambda_0}$ and $\mathfrak{M} \cap \mathfrak{B}_+ = \mathfrak{M}_+$. Put $\mathfrak{p}_+ = \sum_{\gamma \in P_+} CX_\gamma$ and let $\mathfrak{P}_+$ be the subalgebra of $\mathfrak{B}$ generated by $(1, \mathfrak{p}_+)$. Then if $\mathfrak{k}_+ = \mathfrak{k} \cap \mathfrak{g}_+$, $\mathfrak{g}_+ = \mathfrak{p}_+ + \mathfrak{p}_- + \mathfrak{k}_+$ and it follows from Lemma 12 of [5(c)] that $\mathfrak{B}_+ = \mathfrak{P}_- \mathfrak{P}_+ \mathfrak{X}_+$. But $[\mathfrak{p}_+, \mathfrak{k}] \subset \mathfrak{p}_+ \subset \mathfrak{M}_+$ and therefore

$$\mathfrak{P}_+ \mathfrak{X}_+ = \mathfrak{X}_+ + \mathfrak{P}_+ \mathfrak{p}_+ \mathfrak{X}_+ \subset \mathfrak{X}_+ + \mathfrak{M}_+.$$

Hence $\mathfrak{B}_+ = \mathfrak{P}_- \mathfrak{X}_+ + \mathfrak{M}_+$ and $\mathfrak{B}^* = \pi(\mathfrak{B}_+)1^* = \pi(\mathfrak{P}_-)\mathfrak{X}^* = \Gamma(\mathfrak{P}_- \times \mathfrak{X}^*)$. Moreover the mapping $(p \times p' \times x) \to pp'x$ $(p \in \mathfrak{P}_-, p' \in \mathfrak{P}_+, x \in \mathfrak{X}_+)$ defines a linear isomorphism μ of $\mathfrak{P}_- \times \mathfrak{P}_+ \times \mathfrak{X}_+$ onto $\mathfrak{B}_+$ [5(c), Lemma 12]. Therefore

$$\begin{aligned} \mathfrak{M}_+ &= \mathfrak{B}_+ \mathfrak{p}_+ + \mathfrak{B}_+ \mathfrak{Y}_+ = \mathfrak{P}_- (\mathfrak{P}_+ \mathfrak{p}_+) \mathfrak{X}_+ + \mathfrak{P}_- \mathfrak{P}_+ \mathfrak{Y}_+ \\ &= \mu((\mathfrak{P}_- \times (\mathfrak{P}_+ \mathfrak{p}_+) \times \mathfrak{X}_+) + (\mathfrak{P}_- \times \mathfrak{P}_+ \times \mathfrak{Y}_+)). \end{aligned}$$

Hence by considering $\mu^{-1}(\mathfrak{M}_+) \cap \mu^{-1}(\mathfrak{P}_- \mathfrak{X}_+)$ it is obvious that $\mathfrak{P}_- \mathfrak{X}_+ \cap \mathfrak{M}_+ = \mathfrak{P}_- \mathfrak{Y}_+$ and this shows that the kernel of Γ is zero. Therefore Γ is an isomorphism of $\mathfrak{P}_- \times \mathfrak{X}^*$ onto $\mathfrak{B}^*$.

For any linear function Λ on $\mathfrak{h}$ put $n(\Lambda) = \dim \mathfrak{B}_\Lambda{}^*$. We have seen above that $\mathfrak{X}^*$ is irreducible under $\pi(\mathfrak{k})$. Let $\Lambda_0, \Lambda_1, \cdots, \Lambda_r$ be all the distinct weights of the corresponding representation of $\mathfrak{k}$ and let μ_i denote the multiplicity of Λ_i so that $\mu_i = \dim(\mathfrak{B}_{\Lambda_i}{}^* \cap \mathfrak{X}^*)$ $(0 \leqq i \leqq r)$. Then we get the following result from the above lemma.

COROLLARY. *Let* $(\gamma_1, \cdots, \gamma_q)$ *be all the distinct totally positive roots of* $\mathfrak{g}$. *Then if* Λ *is any linear function on* $\mathfrak{h}$,

$$n(\Lambda) = \sum_{i=0}^{r} \mu_i n_i(\Lambda)$$

where $n_i(\Lambda)$ $(0 \leqq i \leqq r)$ *is the number of distinct sequences* $(m, \cdots, m_q)$ *of nonnegative integers such that* $\Lambda = \Lambda_i - (m_1\gamma_1 + \cdots + m_q\gamma_q)$.

Put $\mathfrak{X}_i^* = \mathfrak{B}_{\Lambda_i}^* \cap \mathfrak{X}^*$, $0 \leqq i \leqq r$. Then $\mathfrak{X}^*$ is the direct sum of $\mathfrak{X}_i^*$, $0 \leqq i \leqq r$ and we can choose a base $x_1^*, \cdots, x_d^*$ for $\mathfrak{X}^*$ which is composed of bases for the various $\mathfrak{X}_i^*$. Since $\mathfrak{P}_-$ is the universal enveloping algebra of the abelian Lie algebra $\mathfrak{p}_-$ and since $(X_{-\gamma_1}, \cdots, X_{-\gamma_q})$ is a base for $\mathfrak{p}_-$, it follows that the elements $X_{-\gamma_1}^{m_1} X_{-\gamma_q}^{m_2} \cdots X_{-\gamma_q}^{m_q}$ $(m_1, \cdots, m_q \geqq 0)$ form a base for $\mathfrak{P}_-$. Hence it follows from the above lemma that $\pi(X_{-\gamma_1}^{m_1} \cdots X_{-\gamma_q}^{m_q}) x_j^*$ $(m_1, \cdots, m_q \geqq 0, 1 \leqq j \leqq d)$ is a base for $\mathfrak{B}^*$. Moreover it is obvious that if $x^* \in \mathfrak{X}_i^*$ $(0 \leqq i \leqq r)$

$$\pi(X_{-\gamma_1}^{m_1} X_{-\gamma_2}^{m_2} \cdots X_{-\gamma_q}^{m_q}) x^* \in \mathfrak{B}_\Lambda^*$$

where $\Lambda = \Lambda_i - (m_1\gamma_1 + \cdots + m_q\gamma_q)$. Our assertion now follows immediately from these facts.

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.

REFERENCES.

[1] V. Bargmann, "Irreducible unitary representations of the Lorentz group," *Annals of Mathematics*, vol. 48 (1947), pp. 568-640.

[2] G. Birkhoff and P. M. Whitman, "Representation of Jordan and Lie algebras," *Transactions of the American Mathematical Society*, vol. 65 (1949), pp. 116-136.

[3] E. Cartan, (a) "Les groupes projectifs qui ne lassent invariante aucune multiplicité plane," *Bulletin de la Société Mathématique de France*, vol. 41 (1913), pp. 53-96.

———, (b) "Sur certaines formes riemanniennes remarquables des géométries à groupe fondamental simple," *Annales Scientifiques de l'École Normale Supérieure*, vol. 44 (1927), pp. 345-467.

———, (c) "Groupes simples clos et ouverts et géométrie riemannienne," *Journal de Mathématiques Pures et Appliquées*, 9th Series, vol. 8 (1929), pp. 1-33.

———, (d) "Sur les domaines bornés homogènes de l'espace de n variables complexes," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 11 (1935), pp. 116-162.

———, (e) "La topologie des espaces représentatifs des groupes de Lie," *L'Enseignement mathématique*, vol. 35 (1936), pp. 177-200.

[4] I. M. Gelfand and M. I. Graev, (a) "Unitarnye predstavleniia veshchestvennoi unimoduliarnoi gruppy (osnovnye nevyrozhdennye serii)," *Izvestia Akademiia Nauk SSSR*, vol. 17 (1953), pp. 189-248.

———, (b) "Analog formuly Plansherelia dlia veshchestvennykh poluprostykh grupp Li," *Doklady Akademiia Nauk SSSR*, N. S., vol. 92 (1953), pp. 461-464.

[5] Harish-Chandra, (a) "On representations of Lie algebras," *Annals of Mathematics*, vol. 50 (1949), pp. 900-915.

———, (b) "On some applications of the universal enveloping algebra of a semisimple Lie algebra," *Transactions of the American Mathematical Society*, vol. 70 (1951), pp. 28-96.

———, (c) "Representations of a semisimple Lie group on a Banach space I," *ibid.*, vol. 75 (1953), pp. 185-243.

———, (d) "Representations of semisimple Lie groups. II," *ibid.*, vol. 76 (1954), pp. 26-65.

———, (e) "Plancherel formula for the 2×2 real unimodular group," *Proceedings of the National Academy of Sciences*, vol. 38 (1952), pp. 337-342.

———, (f) "Representations of semisimple Lie groups. V," *ibid.*, vol. 40 (1954), p. 1076.

———, (g) "Representations of semisimple Lie groups. VI," *ibid.*, vol. 40 (1954), p. 1078.

———, (h) "Integrable and square-integrable representations of a semisimple Lie group," *ibid.*, vol. 41 (1955), p. 314.

[6] G. W. Mackey, "Induced representations of locally compact groups. I," *Annals of Mathematics*, vol. 55 (1952), pp. 101-139.

[7] H. Weyl, *The Structure and Representation of Continuous Groups*, The Institute for Advanced Study, Princeton, 1935.

INVARIANTS OF FINITE GROUPS GENERATED BY REFLECTIONS.*

By CLAUDE CHEVALLEY.

---

1. An invertible linear transformation of a finite dimensional vector space V over a field K will be called a *reflection* if it is of order two and leaves a hyperplane pointwise fixed. A group G of linear transformations of V is a *finite reflection group* if it is a finite group generated by reflections. The operations of G extend to automorphisms of the symmetric algebra S of V by the rule $g(P)(x) = P(g^{-1}(x))$, $(P \in S, x \in V)$, and an element $P \in S$ such that $g(P) = P$ for all $g \in G$ is said to be *an invariant of* G. Our main purpose in this note is to prove the theorem:

(A) *Let G be a finite reflection group in a n-dimensional vector-space V over a field K of characteristic zero. Then the K-algebra J of invariants of G is generated by n algebraically independent homogeneous elements (and the unit).*

A vector space A is *graded* by subspaces A^i, (i positive integer), if it is the direct sum of the A^i. The degree d^0P of $P \in A$ is the smallest integer j such that $P \in \sum_{i \leqq j} A^i$; the elements of A^i are the homogeneous elements of degree i. When the A^i are finite dimensional, the Poincaré series of A in the indeterminate t is defined as

$$P_t(A) = \sum_{i \geqq 0} \dim. A^i \cdot t^i.$$

In particular S is graded in the obvious way and $P_t(S) = (1-t)^{-n}$. Let F be the ideal generated by the homogeneous elements of strictly positive degrees in J. Then the grading of S induces a grading of the quotient space S/F. Since F is invariant under G, the operations of G in S induce automorphisms of S/F. We shall also prove:

(B) *Let $I_1, \cdots, I_n$ be a minimal system of homogeneous generators of J and let m_i be the degree of I_i, $(1 \leqq i \leqq n)$. Then*

$$P_t(S/F) = (1-t)^{-n} \cdot \prod_{i=1}^{i=n} (1 - t^{m_i}).$$

---

* Received June 9, 1955.

The product of the m_i is equal to the order of G and to the dimension of S/F. The natural representation of G in S/F is equivalent to the regular representation.

2. Two lemmas. In this paragraph, the characteristic p of the infinite groundfield K is allowed to be $\neq 0$ and G denotes a finite reflection group in V whose order N is prime to p.[1] To any element $P \in S$ we can then associate its average over G:

$$M(P) = 1/N \sum_{g \in G} g(P).$$

LEMMA 1. *Let $U_1, \cdots, U_m$ be invariants of G such that U_1 does not belong to the ideal generated in J by $U_2, \cdots, U_m$. Let P_i, $(1 \leqq i \leqq m)$, be homogeneous elements of S satisfying a relation $\sum_1^m P_i \cdot U_i = 0$. Then $P_1 \in F$.*

If $d^0 P_1 = 0$, then it follows from the assumption and from the relation

$$M(P_1) \cdot U_1 + \cdots + M(P_m) \cdot U_m = 0$$

that $P_1 = M(P_1) = 0$. Assume now $d^0 P_1 > 0$ and the lemma to be true for all relations $\sum_1^m Q_i \cdot U_i = 0$ with homogeneous Q_i and $d^0 Q_1 < d^0 P_1$. Let s be a reflection of G leaving pointwise fixed a hyperplane with equation $L = 0$. Then $s(P_i) - P_i = L \cdot Q_i$, $(Q_i \in S,\ i = 1, \cdots, m)$, and

$$Q_1 \cdot U_1 + \cdots + Q_m \cdot U_m = 0$$

whence, by induction, $Q_1 \in F$ or, otherwise said, $s(P_1) \equiv P_1 \bmod F$; the group G being generated by reflections, we have then $g(P_1) \equiv P_1 \bmod F$ for any $g \in G$, whence $P_1 \equiv M(P_1) \bmod F$; since P_1 is homogeneous of strictly positive degree, the same is true for $M(P_1)$; therefore $M(P_1) \in F$ and $P_1 \in F$.

LEMMA 2. *Assume K to be a perfect field. Let I_i, $(1 \leqq i \leqq m)$, be homogeneous invariants which form an ideal basis of F,[2] with $m_i = d^0 I_i$ prime to p for $i \leqq r$. Then $I_1, \cdots, I_r$ are algebraically independent.*

Let us suppose the lemma to be false and let $H(I_1, \cdots, I_r) = 0$ be a non trivial relation of minimal degree between $I_1, \cdots, I_r$ where $H(y_1, \cdots, y_r)$

[1] In this paper, we are primarily interested in the case $p = 0$, but Lemma 2 will be used in a forthcoming paper of A. Borel, to appear in Jour. Math. Pur. Appl.

[2] This always exists since by the classical theorem for invariants of a finite group, J is a finitely generated K-algebra.

is a polynomial in r letters y_i. We may assume that there exists an integer h such that for any monomial ${y_1}^{k_1} \cdot \cdots \cdot {y_r}^{k_r}$ of H we have

$$k_1 \cdot m_1 + \cdots + k_r \cdot m_r = h.$$

The partial derivatives $\partial H/\partial y_i$ are not all zero, because otherwise (for $p \neq 0$, the only case for which it is not obvious), K being perfect, H would be the p-th power of a polynomial H^*, and $H^*(I_1, \cdots, I_r) = 0$ would be a non trivial relation of strictly smaller degree. Set

$$H_i = \partial H/\partial y_i\,(I_1, \cdots, I_r), \qquad (1 \leqq i \leqq r);$$

then $H_1, \cdots, H_r$ are in J and not all zero; after a possible permutation of indices, we may assume that they belong to the ideal generated in J by the first s of them, but that none of $H_1, \cdots, H_s$ belongs to the ideal generated by the other ones in J. Set

$$H_{s+j} = \sum_{j=1}^{i=s} V_{j,i} H_i.$$

Let x_k, $(1 \leqq k \leqq n)$, be coordinates in V. Since

$$\sum_{i=1}^{i=r} H_i \cdot (\partial I_i/\partial x_k) = 0, \qquad (1 \leqq k \leqq n),$$

we have by Lemma 1

$$\partial I_i/\partial x_k + \sum_{j=1}^{j=r-s} V_{j,i}(\partial I_{s+j}/\partial x_k) \in F, \qquad (1 \leqq i \leqq s;\ 1 \leqq k \leqq n)$$

(the left hand sides are homogeneous in the x_k by the above remark on the monomials of H). Multiplying this relation by x_k and adding the relations thus obtained, we get

$$m_i I_i + \sum_{j=1}^{j=r-s} V_{j,i} m_{s+j} I_{s+j} = \sum_{l=1}^{l=m} A_{i,l} I_l, \qquad (1 \leqq i \leqq s).$$

where the $A_{i,l}$ are forms belonging to the ideal generated by $x_1, \cdots, x_n$. For reasons of homogeneity, we have $A_{i,l} = 0$ if I_l is not of strictly lower degree than I_i; m_i being prime to p for $i \leqq r$, we see that I_i belongs to the ideal generated by the other I_j, which is a contradiction. Thus $I_1, \cdots, I_r$ are algebraically independent.

3. Proofs of Theorems (A) and (B). We assume again the ground-field to be of characteristic zero and denote as in Lemma 2 by $I_1, \cdots, I_m$ homogeneous invariants of G forming an ideal basis of F. By Lemma 2

they are algebraically independent, whence also $m \leqq n$. Using averages over G, it is readily seen by induction on the degree that the unit and the I_i generate J and thus, to finish the proof of (A), there remains to show that $m \geqq n$.

Let $x_1, \cdots, x_n$ be coordinates in V and let $K(x)$ be the field of rational functions in the x_i. It is acted upon in a natural way by G and we denote by L the subfield of elements invariant under G. Then $K(x)$ is a Galois extension of L, with Galois group G and L has also transcendence degree n over K. On the other hand, G being *finite*, every invariant in $K(x)$ is classically the quotient of two invariant polynomials; thus $L = K(J)$ is generated by the I_i, and $m \geqq n$.

LEMMA 3. *Let $P_1, \cdots, P_s$ be homogeneous elements of S whose residue classes* mod F *are linearly independent over K in S/F. Then $P_1, \cdots, P_s$ are linearly independent over $K(J)$.*

Let $V_1 \cdot P_1 + \cdots + V_s \cdot P_s = 0$ be a relation with $V_i \varepsilon K(J)$, $(1 \leqq i \leqq s)$. We have to prove that $V_i = 0$ for all i and it is enough to consider the case where the V_i are homogeneous elements of J such that $d^0 V_i + d^0 P_i$ is equal to a constant h independent of i.

By the degree of the monomial $I_1^{k_1} \cdots \cdots I_n^{k_n}$ we mean its degree as element of S, i.e. $k_1 m_1 + \cdots + k_n m_n$. Let S_j, $(j = 1, 2, \cdots)$, be the different monomials in the I_i arranged by increasing degrees, with $S_1 = 1$. We have

$$V_i = \sum_{j \geqq 0} k_{ij} S_j, \qquad (k_{ij} \varepsilon K, k_{ij} = 0 \text{ for } d^0 V_i \neq d^0 S_j, \ (1 \leqq i \leqq n)),$$

and our relation may be written

$$\sum_{j \geqq 0} W_j \cdot S_j = 0, \qquad (W_j = \sum_{i=1}^{i=s} k_{ij} P_i),$$

where W_j is homogeneous, of degree equal to $h - d^0 S_j$. Assume that $k_{ij} = 0$ for $1 \leqq i \leqq s$ and $j < t$. Since by Theorem A the monomial S_t does not belong to the ideal generated in J by the S_j with $j > t$, we have by Lemma 1 $W_t \varepsilon F$ and the hypothesis gives then $k_{it} = 0$ for $i = 1, \cdots, s$. This proves by induction on j that $k_{ij} = 0$ for all i, j, and the lemma.

We now come to the proof of (B). The field $K(x)$ being a normal extension of $K(J)$ with Galois group G, has degree N over $K(J)$, hence the dimension of S/F over K is finite. Let $A_1, \cdots, A_q$ be homogeneous polynomials whose residue classes mod F form a basis of S/F. By induction on the degree we see that every $P \varepsilon S$ may be expressed as linear combination

of the A_i with coefficients in J, and this expression is unique in view of Lemma 3. Hence

$$P_t(S) = P_t(S/F) \cdot P_t(J);$$

but $P_t(S) = (1-t)^{-n}$ and Theorem A gives $P_t(J) = \prod_1^n (1-t^{m_i})^{-1}$, whence the first assertion of (B). We may also write

$$P_t(S/F) = \prod_{i=1}^{i=n} (1 + t + t^2 + \cdots + t^{m_i-1})$$

and, setting $t=1$, we get $\dim. S/F = m_1 \cdot \cdots \cdot m_n$. Since every element of $K(x)$ may be written as the quotient of a polynomial by an invariant polynomial, it also follows from the above and Lemma 3 that the A_i form a basis of $K(x)$ over $K(J)$, whence $N = \dim. S/F$.

We have for $g \in G$

$$g(A_i) = \sum_{j=1}^{j=N} a_{ij}(g) A_j, \qquad (i = 1, \cdots, N),$$

where the $a_{ij}(g)$ are homogeneous elements of J and where $a_{ii}(g) \in K$ by homogeneity. The matrices $(a_{ij}(g))$ describe the natural representation of G in $K(x)$, considered as vector space over $K(J)$. If we reduce the coefficients mod F we get the natural representation of G in S/F, considered as vector space over K; this reduction does not affect the diagonal coefficients, hence both representations have the same character and are equivalent. But G is the Galois group of the normal extension $K(x)$ of $K(J)$, so that the former representation is equivalent to the regular representation, which proves the last statement of (B).

COLUMBIA UNIVERSITY.

ON THE ARITHMETIC STRUCTURE OF A CLASS OF COMMUTATIVE SEMI-GROUPS.*

By LEE M. SONNEBORN.

1. Introduction. Let S be a commutative semi-group [Jacobson [3], Chapter IV] in which both chain conditions hold for non-associated elements, and consider a multiplicative diophantine system [Skolem [6], Chapter IV] of the form

$$x_1^{a_{i1}}x_2^{a_{i2}}\cdots x_k^{a_{ik}} = y_1^{b_{i1}}y_2^{b_{i2}}\cdots y_h^{b_{ih}} \qquad (i=1,2,\cdots,m) \tag{1}$$

where the a_{ij} and b_{ij} are non-negative integers and the x_i and y_i are indeterminates over S.

It was shown by E. T. Bell [1] that if S is the semi-group of positive integers, the solution of (1) may be reduced to the special case

$$x_1x_2 = y_1y_2, \qquad \text{for which} \tag{2}$$

$$x_1 = u_1v_1,\ x_2 = u_2v_2,\ y_1 = u_1u_2,\ y_2 = v_1v_2. \tag{3}$$

We investigate here what restrictions are imposed on S if, whenever (2) holds in S, elements u_1, u_2, v_1, v_2 exist in S such that (3) is satisfied. Such a system S is said by Birkhoff [2] to satisfy the Riesz Refinement Postulate after F. Riesz [5]. We shall prove that upon identifying associated elements, S becomes the sub-direct product of a distributive lattice, and a Gaussian semi-group [Jacobson, loc. cit.].

2. We say a "divides, or does not divide," b ($a|b$ or $a\nmid b$) if there exists, or does not exist an element, c, of S such that $ac=b$. If a divides b and b divides a, then a and b are called associates ($a\sim b$). Otherwise $a\nsim b$. If S contains an element e_a such that $e_aa=a$, then e_a is called an identity for a. If $e|x$ for all x in S, then e is called a unit. It is not, in general, true that two associate elements differ by a unit, as is shown by the system

	1	a	b
1	1	a	b
a	a	a	b
b	b	b	a

in which $a\sim b$, but 1 is the only unit.

* Received May 18, 1955.

We consider now an associative, commutative semi-group S in which

(1) *Every ascending or descending chain of non-associated divisors of any element, a, of S is of finite length.*

(2) *$ab = cd$ if and only if there exist elements x, y, z, w in S such that $a = xy$, $b = zw$, $c = xz$, and $d = yw$.*

On applying (2) to the equation $aa = aa$, we find that

$$a = x_1y_1 = z_1w_1 = x_1z_1 = y_1w_1.$$

If $x_1 \neq a$, then $x_1x_1 = x_1x_1$ implies

$$x_1 = x_2y_2 = z_2w_2 = x_2z_2 = y_2w_2.$$

Continuing in this fashion we obtain a chain of divisors of $a, x_1, x_2 \cdots x_i \cdots$, in which each element divides all the preceding elements. By condition (1) above, only a finite number of the x_i can be non-associates. Hence $x_n \sim x_{n'}$ for some n and n'. But then $a = kx_n = lx_{n'}$ for some k and l, and $x_n = mx_{n'}$ and $x_{n'} = px_n$. Hence $mpa = kmpx_n = kmx_{n'} = kx_n = a$, so mp is an identity for a. But a was arbitrary; hence we have proved

(3) *For each element a of S, there exists an e_a, called an identity for a, such that $e_aa = a$.*

Condition (3) implies that the relation $\sim$ is reflexive. *Hence $\sim$ is an equivalence relation*; for its symmetry and transitivity are obvious from the definition.

THEOREM 1. *The system S', whose elements are the equivalence classes of S under $\sim$, is an associative commutative semi-group in which*

(1′) *Any ascending or descending chain of proper divisors of an element of S' is of finite length*

(2′) *$ab \sim cd$ if and only if $a \sim xy$, $xz \sim st \mid c$, $c \sim su$, and*
$$b \sim zw,\ yw \sim uv \mid d,\ d \sim tv.$$

(3′) *For any a in S', there is an e_a in S' such that $e_aa = a$.*

Proof. Since $\sim$ is defined by division, $a \sim a'$ and $b \sim b'$ imply $a = ka'$, $a' = la$, $b = mb'$, $b' = nb$ for some k, l, m, and n, and hence $ab = kma'b'$ and $a'b' = lnab$ which gives $ab \sim a'b'$. Thus multiplication is well defined in S'. Properties (1′) and (3′) then follow immediately from conditions (1) and (3) in S. The proof of (2′) rests on two lemmas.

Consider the co-ideal $\{a\}$, [MacKenzie [4]], consisting of all divisors of the element a. Clearly $\{a\} \subseteq \{b\}$ if and only if $a|b$, and $\{a\} = \{b\}$ if and only if $a \sim b$.

LEMMA 1. *$\{a\}\{b\} = \{ab\}$ where $\{a\}\{b\}$ is the set of all elements a_1b_1 with $a_1 \varepsilon \{a\}$, and $b_1 \varepsilon \{b\}$.*

Proof. Trivially $\{a\}\{b\} \subseteq \{ab\}$. Let $c \varepsilon \{ab\}$. Then $c|ab$ and therefore $ab = cd$ for some d. But then, by (1) $a = xy$, $b = zw$, $c = xz$, $d = yw$, where c has been displayed as xz and $x|a$ and $z|b$. Let $a_1 = x$, $b_1 = z$. Thus $c \varepsilon \{a\}\{b\}$.

The following two corollaries, amounting to little more than restatement, will be useful later:

COROLLARY 1. *If $\{c\} \subseteq \{ab\}$, then $c = c_1c_2$ where $c_1 \varepsilon \{a\}$, and $c_2 \varepsilon \{b\}$.*

COROLLARY 2. *If $\{cd\} = \{ab\}$, then $ab = a_1b_1$ where $a_1 \varepsilon \{c\}$ and $b_1 \varepsilon \{d\}$.*

LEMMA 2. *If $\{ab\} = \{cd\}$, then there exist a', b', c', and d' such that $ab = a'b'$, $cd = c'd'$, and $\{a'\} = \{c'\} \subseteq \{c\}$ and $\{b'\} = \{d'\} \subseteq \{d\}$.*

Proof. We form a chain which is forced to terminate by condition (1) in S. By Corollary 2, we choose a_1 and b_1 such that $a_1b_1 = ab$ and $\{a_1\} \subseteq \{c\}$ and $\{b_1\} \subseteq \{d\}$. Then from $\{a_1b_1\} = \{cd\}$ we choose c_1 and d_1 so that $c_1d_1 = cd$ and $\{c_1\} \subseteq \{a_1\} \subseteq \{c\}$ and $\{d_1\} \subseteq \{b_1\} \subseteq \{d\}$. Continuing in this fashion, we obtain a_2 and b_2 such that $ab = a_1b_1 = a_2b_2$ and

$$\{a_2\} \subseteq \{c_1\} \subseteq \{a_1\} \subseteq \{c\}, \text{ and } \{b_2\} \subseteq \{d_1\} \subseteq \{b_1\} \subseteq \{d\}$$

and, eventually,

$$ab = a_1b_1 = a_2b_2 = a_3b_3 = \cdots = a_nb_n = \cdots$$

and

$$cd = c_1d_1 = c_2d_2 = c_3d_3 = \cdots = c_nd_n = \cdots,$$

in which

$$\{a_n\} \subseteq \{c_{n-1}\} \subseteq \{a_{n-1}\} \subseteq \{c_{n-2}\} \subseteq \cdots \subseteq \{c\},$$

and

$$\{b_n\} \subseteq \{d_{n-1}\} \subseteq \{b_{n-1}\} \subseteq \{d_{n-2}\} \subseteq \cdots \subseteq \{d\}.$$

But now, by virtue of condition (1) in S, there is an m such that $\{a_m\} = \{c_m\}$ and $\{b_m\} = \{d_m\}$. Let $a_m = a'$, $b_m = b'$, $c_m = c'$ and $d_m = d'$ and the lemma is proved.

We may now conclude the proof of (2′) and hence of Theorem 1 as follows: We have $\{ab\} = \{cd\}$ if and only if $ab \sim cd$. Then by Lemma 2,

$ab \sim cd$ implies $ab = a'b'$, and $c'd' = cd$ where $a' \sim c'$, $b' \sim d'$ and $a' | c$, and $b' | d$. But by condition (2) in S, we have

$$a = xy,\ a' = xz \sim st = c' | c,\ c = su;$$
$$b = zw,\ b' = yw \sim uv = d' | d,\ d = tv.$$

Clearly equality implies equivalence, so all equalities in the above may be replaced by equivalences, thus completing the proof of Theorem 1.

3. In the remainder of the paper the system S' is considered where equivalence in S is replaced by equality in S'. S' is then an associative, commutative semi-group in which

(1′) *Any ascending or descending chain of proper divisors of any element of S' is of finite length.*

(2′) *$ab = cd$ if and only if there exist x, y, z, w, s, t, u, and v such that*

$$a = xy,\ xz = st | c,\ c = su;$$
$$b = zw,\ yw = uv | d,\ d = tv.$$

(3′) *For any a in S', there is an identity for a, and*

(4′) *If $a | b$ and $b | a$, then $a = b$.*

THEOREM 2. *There is an element 1 in S' such that $1a = a$ for all a in S.*

Proof. Clearly any divisor of an identity for a is an identity for a, by virtue of condition (4′). Also any two identities for the same element have a common divisor, for if $e_a a = e_a' a = a$, we have from (2′) $e_a = xy$, $xz = st | e_a'$, $e_a' = su$ etc., where $x | e_a$ and $x | e_a'$. Now, if $e_a a = a$ and $e_b b = b$, e_a and e_b have a common divisor which is an identity for a and b, since $e_a ab = e_b ab = ab$. Consider now a maximal chain of divisors of e_a where $e_a a = a$. It has a smallest element e_{0_a} which has the property that it divides all elements of the chain. If a second chain is picked, its smallest element e_{0_a}' must equal e_{0_a} since the two minimal elements must have a common divisor. Hence e_{0_a} depends only on a, not on the chain selected. For any other element b there is similarly an identity e_{0_b}. But then, by the above e_{0_a} and e_{0_b} have a common divisor which must be equal to both by minimality. Hence $e_{0_a} x = x$ for all x and e_{0_a} is the element 1 with properties stated in Theorem 2.

4. We define an element p, to be irreducible if the only divisors of p are identities for p and p itself. Clearly any element of S' can be written as a finite product of irreducibles in at least one way because of the descending chain condition.

THEOREM 3. *The set of idempotent elements of S' is identical with the set of elements which are identities.*

Proof (a) *Any non-idempotent irreducible has all its powers distinct.* For suppose p is irreducible and $p \neq p^2$ and $p^n = p^{n+m}$ for some n and m. Then $p^n \mid p^{n+1} \mid p^{m+n} = p^n$ implies by (4') that $p^n = p^{n+1}$. We pick the smallest n for which this is true ($n \geqq 2$) and eventually arrive at the contradiction that $p^n = p^{n+1}$ implies $p^{n-1} = p^n$. Now $p^{n+1} = p^n$ is the same as $p \cdot p^n = p \cdot p^{n-1}$ where the only divisors of p are p and identities for p. But

$$p = xy,\ xz = st \mid p,\ su = p;$$

$$p^n = zw,\ yw = uv \mid p^{n-1},\ tv = p^{n-1}.$$

Clearly, x or y, and s or u, must be p since p is non-idempotent and irreducible.

Case I. $x = s = p$; $y = e_p'$, $u = e_p$, where e_p and e_p' are identities for p. Then z and t must be identities for p, say e_p'' and e_p'''. Then $p^n = zw = e_p''w$ and $e_p'w = yw \mid p^{n-1}$. But then $e_p''e_p'w \mid e_p''p^{n-1} = p^{n-1}$ and $e_p'e_p''w = e_p'p^n = p^n$ from which $p^n \mid p^{n-1}$, or $p^n = p^{n-1}$, the desired contradiction.

Case II. $y = s = p$; $x = e_p$, $u = e_p'$, $t = e_p''$, $z = p$. But then, similarly $p^n = zw = pw$ and $yw = pw \mid p^{n-1}$.

Case III. $x = u = p$; $y = e_p$, $z = e_p'$, $s = e_p''$ and $t = p$. Then $p^n = zw = e_p'w$, $yw = e_pw \mid p^{n-1}$. But $e_pp^n = p^n = e_pe_p'w$ and

$$e_p'yw = e_p'e_pw \mid e_p'p^{n-1} = p^{n-1} \text{ or } p^n = e_pe_p'w \mid p^{n-1}.$$

Case IV. $y = u = p$; $x = e_p$, $s = e_p'$. Two cases are possible. First if $z = p$, then $p^n = pw$, $yw = pw \mid p^{n-1}$, or $p^n = p^{n-1}$. Otherwise $z = e_p''$ in which case $p^n = zw = e_p''w$ and $pw \mid p^{n-1}$ whence $p \cdot p^n = p \cdot e_p''w = pw \mid p^{n-1}$ from which $p^{n+1} \mid p^{n-1}$ which is false and completes the proof of (a).

(b) *Any divisor of an idempotent is idempotent.* For let a be idempotent and $a = pb$ where p is irreducible. Then $a^n = p^nb^n = a$. Hence $p^n \mid a$ for all n. But the p^n are an infinite ascending chain of divisors of a unless p is idempotent. Hence p is idempotent. But then any irreducible divisor of an idempotent is idempotent. But any divisor of an idempotent is a product of irreducibles and is hence idempotent.

(c) *Any identity is idempotent.* Consider an element e_a with the property that $e_a a = a$. Let p, an irreducible, divide e_a. Then $p^n | a$ for all a and hence p is idempotent. But e_a is the product of a finite number of such idempotents, and is thus idempotent. But clearly any idempotent is an identity (for itself) and so the set of identities is the same as the set of idempotents completing the proof of Theorem 3.

THEOREM 4. *The identities of S' form a distributive lattice.*

Proof. Let I be the set of idempotents partially ordered by $x \leqq y$ if $y | x$, and consider the product ab with a and b in I. Then ab is in I and $ab \leqq a$, $ab \leqq b$. Assume x in I and $x \leqq a$, $x \leqq b$. Then $x = ka$ and $ax = aak = ak = x$ and $abx = x$. Hence $ab | x$ and $x \leqq ab$. Thus $ab = a \cap b$, and I is a lattice, for it is a partially ordered set with a unit element 1, the descending chain condition, and is closed under finite intersection. I is distributive for if $a \geqq b \cap c$, there exist b_1 and c_1 such that $a = b_1 \cap c_1$ with $b_1 \geqq b$ and $c_1 \geqq c$. This is easily shown to be equivalent to distributivity in a lattice, and, rewritten in terms of division, is precisely Corollary 1 of Section 2. Now, $ab = cd$ if and only if $a = xy$, $b = zw$, $c = xz$, $d = yw$, for distributivity in a lattice is also equivalent to $a \cap b = c \cap d$ if and only if $a = (a \cup c) \cap (a \cup d)$, $b = (b \cup c) \cap (b \cup d)$, $c = (a \cup c) \cap (b \cup c)$ and $d = (a \cup d) \cap (b \cup d)$. Hence *condition (2) of S holds among the identities of S'.*

5. THEOREM 5. *If an element of S' is factored into irreducibles in two ways, the non-idempotents and their powers are identical in the two factorizations.*

Proof. Let p, p_1, p_2 denote non-idempotent irreducibles.

(a) If $p | ab$ then $p | a$ or $p | b$. For let $pk = ab$. Then by condition (2′) on S',

$$p = xy \quad xz = st | a,$$

where x or $y = p$ and $x | a, y | b$.

$$k = zw \quad yw = uv | b,$$

(b) Any non-idempotent irreducible has no non-idempotent divisors except itself.

(c) In two factorizations of an element into irreducibles the distinct non-idempotents occurring are the same. For, let p occur in one factorization, then by (a) it must divide a p' in the second factorization and by (b) it must be the same element.

(d) No non-idempotent irreducible divides any power of any other non-idempotent. For let $p_1 \neq p_2$ and $p_1 | p_2^{\alpha}$ where α is minimal with this property. Then $p_1 | p_2^{\alpha-1} p_2$. But by (a), $p_1 | p_2$ or $p_1 | p_2^{\alpha-1}$ both of which are false.

(e) If $p^n | ab$ and $p \nmid a$ then $p^n | b$. We prove this by induction on n. By (a) this is true for $n = 1$. Let m be the smallest integer for which $p^m | ab$, $p \nmid a$ and $p^m \nmid b$. By the induction hypothesis, $p^{m-1} | ab$ and $p \nmid a$ implies $p^{m-1} | b$. From $kp^m = ab$, we have by (2′)

$$p^m = xy, xz = st | a, su = a;$$

$$k = zw, yw = uv | b, tv = b.$$

But $p \nmid a$ and hence x, z, s, t and u are not divisible by p. But then $p^m = xy$ where $p \nmid x$ and hence $p^{m-1} | y$. Hence $p \cdot p^{m-1} = xy$ from which

$$p^{m-1} = x'y', x'z' = s't' | x, s'u' = x;$$

$$p = z'w', y'w' = u'v' | y, t'v' = y.$$

But p cannot divide x, or hence x', z', t' or u'. But if $p \nmid x'$, then $p^{m-1} | y'$ and if $p \nmid z'$, $p = w'$. Thus $p \cdot p^{m-1} = p^m | w'y' | y$ and $y | b$, so $p^m | b$ contrary to assumption. Now, if an element a of S' has two factorizations

$$a = e_a p_1^{\alpha_1} \cdots p_r^{\alpha_r} = e_\beta p_1^{\beta_1} \cdots p_r^{\beta_r}$$

in which e_α and e_β are the idempotent parts, $\alpha_i = \beta_i$ for all i since $p_i^{\alpha_i} | k p_i^{\beta_i}$ where $p_i \nmid k$. This completes the proof of Theorem 5.

The set of idempotents which divide a, an element of S', has a maximal element e_a. Say it has a second e_a'. Then $e_a a = e_a' a = a = e_a e_a' a$. But since e_a and e_a' are maximal they are equal. Hence we may factor the element a uniquely into e_a, its maximal idempotent divisor, and powers of irreducible non-idempotents. Let $a = e_a p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, and $b = e_b p_1^{\beta_1} \cdots p_r^{\beta_r}$ (α_i and $\beta_i \geqq 0$) be factorizations in this unique fashion. Then

$$ab = e_a e_b p_1^{\alpha_1+\beta_1} \cdots p_r^{\alpha_r+\beta_r}$$

is the factorization of their product. For clearly $e_a e_b | ab$, and $e_{ab} | ab$ implies, by Corollary 1 of Section 2, that $e_{ab} = e_a' e_b'$ such that $e_a' | e_a$ and $e_b' | e_b$. But this in turn implies that $e_{ab} = e_a' e_b' | e_a e_b | e_{ab}$. Hence $e_a e_b = e_{ab}$.

THEOREM 6. *$ab = cd$ if and only if there exist x, y, z, and w such that $a = xy$, $b = zw$, $c = xz$ and $d = yw$.*

Proof. $a = e_a p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, $b = e_b p_1^{\beta_1} \cdots p_r^{\beta_r}$, $c = e_c p_1^{\gamma_1} \cdots p_r^{\gamma_r}$ and

$d = e_d p_1^{\delta_1} \cdots p_r^{\delta_r}$. By the above $ab = cd$ if and only if $e_a e_b = e_c e_d$ and $\alpha_i + \beta_i = \gamma_i + \delta_i$. From the remark on distributive lattices $e_a = x_1 y_1$, $e_b = z_1 w_1$, $e_c = x_1 z_1$ and $e_d = y_1 w_1$, and from the truth in integers (or arbitrary Gaussian semi-groups), the truth for irreducibles is immediate. The product of the separate parts clearly suffices.

THEOREM 7. *The system S' is the sub-direct product of a distributive lattice and a Gaussian semi-group.*

Proof. Replace $a = e_a p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ by the pair $(e_a, p_1^{\alpha_1} \cdots p_r^{\alpha_r})$. Then $ab = (e_a \cap e_b,\ p_1^{\alpha_1+\beta_1} \cdots p_r^{\alpha_r+\beta_r})$. However, the product is not necessarily direct. For consider the set $R = \{(e_0, 1), (e_1, 1), (e_1, p^n)\ (n = 1, 2, 3, \cdots)\}$ with $e_0 \cap e_1 = e_1$. All the postulates are satisfied, but R is not a direct product, for the $\{(e_0, p^n)\}$ are missing.

The question of whether or not any diophantine system may be reduced as in Bell [1] to the special case considered here will be treated elsewhere.

The author wishes to thank Professor Morgan Ward who suggested the problem, and whose helpful suggestions and encouragement are greatly appreciated.

CALIFORNIA INSTITUTE OF TECHNOLOGY.

REFERENCES.

[1] E. T. Bell, "Reciprocal arrays and diophantine analysis," *American Journal of Mathematics*, vol. 55 (1933), pp. 50-66.

[2] G. Birkhoff, "Lattice-ordered groups," *Annals of Mathematics*, vol. 43 (1942), p. 328.

[3] N. Jacobson, "Lectures in Abstract Algebra," vol. 1, Chap. IV, Van Nostrand, New York, 1951.

[4] R. E. MacKenzie, "Commutative semi-groups," *Duke Mathematical Journal*, vol. 21 (1954), pp. 471-477.

[5] F. Riesz, "Sur quelques notions fondamentales dans la théorie générale des operations lineaires," *Annals of Mathematics*, vol. 41 (1940), p. 175.

[6] Th. Skolem, "Diophantische Gleichungen," *Ergebnisse der Mathematik und ihrer Grenzgebiete*, vol. 5 (1938), Chap. IV.

ON A PROBLEM OF POINCARÉ CONCERNING RICCATI'S EQUATION.*

By PHILIP HARTMAN and AUREL WINTNER.

---

1. For large positive t, consider the differential equation

$$x'' + f(t)x = 0 \tag{1}$$

in which $f(t)$ is a real-valued, continuous function having the property that

$$\omega = \lim_{t\to\infty} f(t) \text{ exists} \tag{2}$$

(as a finite limit). Then equation (1) can be considered to be a "small" perturbation of the equation with constant coefficients

$$x'' + \omega x = 0. \tag{3}$$

This prototype of (1) has the pair of solutions $x = e^{\pm\lambda t}$, $x = e^{\pm i\lambda t}$ or $x = 1, t$ according as $\omega = -\lambda^2 < 0$, $\omega = \lambda^2 > 0$ or $\omega = 0$. Correspondingly, an assertion of Poincaré [9], pp. 227-236, implies that (1), in the case (2), has solutions $x = z(t)$ with a logarithmic derivative

$$w = z'/z \tag{4}$$

satisfying

$$w^2(t) \to -\omega \text{ as } t\to\infty. \tag{5}$$

But the indications given by Poincaré are incomplete, and the purpose of this paper is to examine the validity of the assertion (when $\omega = f(\infty) \geqq 0$).

What is at stake is the existence, and not the evaluation, of a limit, $w(\infty)$, of (4) as $t\to\infty$. For if it is assumed that $w(\infty)$ exists, then $w^2(\infty) = -\omega$ (so that (5) holds). In order to see this, note that (4) satisfies the Riccati equation

$$w' + w^2 + f(t) = 0$$

(at those t-values at which $z \neq 0$). Hence the assumptions that $f(\infty) = \omega$ and $w(\infty)$ exist imply that $w'(t)$ tends to $-w^2(\infty) - \omega$ as $t\to\infty$. But the existence of $w(\infty)$ means that $w'(t)$ is improperly integrable at $t = \infty$, and so the limit of $w'(t)$ at $t = \infty$ must be 0, which proves that $w^2(\infty) = -\omega$.

---

* Received May 5, 1955.

Poincaré's statement is correct if $f(\infty) < 0$ and, in this case,* every non-trivial ($\not\equiv 0$), real- or complex-valued solution $x = z(t)$ of (1) satisfies (4)-(5); see Perron [6]. When

$$\omega = 0, \tag{6}$$

then (1) can be non-oscillatory or oscillatory, where, in the terminology of A. Kneser, (1) is called non-oscillatory if one (hence, by Sturm's separation theorem, every) real-valued solution ($\not\equiv 0$) of (1) does not vanish for large t. In the non-oscillatory case (2), (6) of (1), it is known that every (real- or complex-valued) solution $x = z(t)$ ($\not\equiv 0$) of (1) satisfies (4), (5); see [2], p. 391, Theorem B.

If (1) is oscillatory and $x = z(t)$ is a real-valued solution, then the ratio (4) is undefined at the zeros of $z(t)$. On the other hand, if $x = x(t)$ and $x = y(t)$ are two real-valued, linearly independent solutions of (1) (which, therefore, do not vanish simultaneously), then the complex-valued solution

$$z = x(t) + iy(t) \tag{7}$$

of (1) has no zeros (so that (4) is defined for all t). Perron [7] has shown that, nevertheless, Poincaré's assertion in the case (6) is false; that is, that there exist differential equations (1) which have coefficient functions satisfying (2), (6) (and which are necessarily oscillatory) but are such that no (complex-valued) solution $x = z(t)$ satisfies (4), (5). Strictly speaking, Poincaré's paper assumes that $f(t)$ is a rational function, but this is of no avail (in this connection, cf. the positive results of Perron [6] as well as his negative observations in [6], pp. 269-270).

The question of the validity of Poincaré's assertion in the case $\omega = f(\infty) > 0$, say

$$\omega = 1, \tag{8}$$

does not appear to have been treated in the literature. The chances for the truth of the assertion in this case seem to be higher than they are in the case (6). In fact, the prototypes (3) in the cases (6) and (8) are $x'' = 0$ and $x'' + x = 0$, and the former has secular solutions (i.e., it leads to a

---

* In view of the superposition principle, the assertion is implied by the statement that if $f(t)$ is a real-valued continuous function and if (1) has a real-valued solution $x(t)$ for which neither $x(t)$ nor $x'(t)$ vanishes for large t, then (2) implies (5) for the case $z(t) = x(t)$ of (4).

A simple proof follows by applying l'Hôpital's rule to the function x'^2/x^2. In fact, the ratio of the derivatives of x'^2 and x^2 is $2x'x''/2xx' = x''/x$, which is $-f$, by (1), and so the assertion follows from (2).

multiple elementary divisor), whereas the roots of the characteristic equation of the latter are distinct ($\lambda = \pm i$), and even the title of Perron's paper [7] emphasizes the multiple character of the roots in his case (6). It turns out, however, that Poincaré's assertion is false in the case (8) also. In fact, if

$$f(t) = 1 + (\sin t)/(\tfrac{1}{2}t), \tag{9}$$

then (2) and (8) are satisfied (and $f(t)$, instead of being just continuous, is an entire function) but (1) fails to possess any solution $x = z(t)$ for which Riccati's quotient z'/z tends to a limit as $t \to \infty$.

In order to see this, let

$$r = r(t) = |z(t)| = (x^2 + y^2)^{\frac{1}{2}}. \tag{10}$$

Then

$$rr' = xx' + yy'; \tag{11}$$

so that, from (4),

$$w = r'/r + ic/r^2, \tag{12}$$

where c is the Wronskian constant,

$$xy' - x'y = c\ (\neq 0). \tag{13}$$

It is clear from (12) that (5) cannot hold if $r = r(t)$ is unbounded, as $t \to \infty$. This will be the case if (1) has some unbounded (real-valued) solution; that is, if (1) has some solution satisfying

$$\limsup_{t \to \infty} |x(t)| = \infty. \tag{14}$$

But Perron [8] has shown, in another connection, that (2) and (8) are compatible with (14); in fact, (9) is a simplified version of his example (cf. [10], pp. 386-387). Hence, the case (9) of (1) has no solution satisfying (5), where $\omega = 1$.

The considerations leading to these remarks will be considerably refined below.

2. The following conventions will be used: The assumption that $f(t)$ is real-valued and continuous for large t will not be repeated below. $x(t)$ and $y(t)$ denote real-valued, linearly independent solutions of (1) and $c\ (\neq 0)$ denotes their Wronskian constant (13). A solution $z = z(t)$ of (1) will be called of type (7), (13) if it is the linear combination (7) of real-valued solutions satisfying (13) with a $c \neq 0$.

(I) *Let* (1) *possess a solution* $z=z(t)$ *of type* (7), (13) *having a Riccati quotient* (4) *which satisfies* (5), *where* ω *is (finite and) real. Then*

$$(15_\omega) \qquad \underset{0<s<\infty}{\text{l. u. b.}} \left| \int_t^{t+s} (f(u)-\omega)\,du \right| /(1+s) \to 0 \text{ as } t\to\infty.$$

This shows the extent to which (2) can be altered when (4), (5) hold. The theorem is known in the cases that $\omega<0$ in (5) ([1], pp. 570-571) and in the case that (1) is non-oscillatory and $\omega=0$ (see [2], p. 391).

The contrast between the cases $\omega=0$ and $\omega>0$ is shown by the following theorems:

(II_0) *Let* $z=z(t)$ *be a solution of* (1) *of type* (7), (13) *for which the Riccati ratio* (4) *satisfies*

$$(16_0) \qquad w(t)\to 0 \text{ as } t\to\infty.$$

Then every solution of type (7), (13) *satisfies* (16_0).

If (16_0) is replaced by (5), where $\omega>0$, say by

$$(16_+) \qquad w(t)\to i \text{ as } t\to\infty,$$

then there is a uniqueness theorem.

(II_1) *Let* $z=z(t)$ *be a solution of* (1) *of type* (7), (13) *for which* (4) *satisfies* (16_+). *Then the solution* $z=z(t)$ *is unique up to a constant factor.*

If (16_-) means the relation which results if i is replaced by $-i$ in (16_+), then $z=z(t)$ satisfies (16_+) if and only if its complex conjugate $\bar{z}(t)$, which is also a solution of type (7), (13), satisfies (16_-).

(III_0) *Let* $z=z(t)$ *be a solution of* (1) *of type* (7), (13). *Then the conditions* (15_0) *and*

$$(17) \qquad |z(t)|\to\infty \text{ as } t\to\infty$$

are necessary and sufficient for (16_0).

The analogue for ($16_\pm$) is as follows:

(III_1) *Let* $z=z(t)$ *be a solution of* (1) *of type* (7), (13). *Then the condition* (15_1) *and the condition that*

$$(18) \qquad \lim_{t\to\infty} |z(t)| \text{ exists}$$

(as a finite limit) are necessary and sufficient for ($16_\pm$).

In this theorem, (18) can be replaced by the condition that

$$\lim_{t\to\infty} |z'(t)| \text{ exists.} \tag{19}$$

(IV_1) *Let* $z=z(t)$ *be a solution of* (1) *of type* (7), (13). *Then* (15_1) *and* (19) *are necessary and sufficient for* ($16_\pm$).

It will be seen from the proofs that if (15_1) is assumed, then the necessary and sufficient condition (18) and/or (19) for ($16_\pm$) can be replaced by the condition that $\lim_{t\to\infty}(|z(t)|^2+|z'(t)|^2)$ exists and is $2|c|$.

For the case $\omega=1$ in (5), there results the following existence theorem:

(V_1) *If* (15_1) *holds, then necessary and sufficient for the existence of a solution* $z=z(t)$ *of* (1) *of type* (7), (13) *satisfying* (16_+) *is that*

$$\lim_{t\to\infty} |x+ix'| \text{ exists} \tag{20}$$

for every (real-valued) solution $x=x(t)$ *of* (1).

It may be remarked that (20), as a sufficient condition, cannot be relaxed to the assumption that

$$|x+ix'|=O(1) \text{ as } t\to\infty \tag{21}$$

holds for every (real-valued) solution $x=x(t)$ of (1); cf. Section 9 below. It will, however, remain undecided whether or not the sufficient condition (20) can be relaxed to the assumption that (20) holds for *some* (instead of *every*) real-valued solution $x=x(t)\not\equiv 0$.

(VI_1) *If* (1) *has a solution* $z=z(t)$ *of type* (7), (13) *satisfying* (16_+) *and if* $h=h(t)$ *is a continuous function for large positive* t *satisfying* $\int^\infty |h(t)|\,dt<\infty$, *then* $x''+(f(t)+h(t))x=0$ *has a solution* $z=z(t)$ *of type* (7), (13) *satisfying* (16_+).

In fact, since the solutions of (1) are bounded, as $t\to\infty$, by (III_1) and/or (IV_1), it follows from [11], p. 261, that to any solution $x=x(t)$ of (1) there corresponds a (unique) solution $x=x_0(t)$ of the last equation satisfying, as $t\to\infty$,

$$x_0(t)=x(t)+o(1), \qquad x_0'(t)=x'(t)+o(1).$$

Thus, (VI_1) is a consequence of (V_1) (for instance).

3. Proof of (I). Let $z=z(t)$ be a solution of (1) for which the Riccati ratio (4) satisfies (5). Since (I) is known when $\omega<0$, it will be

supposed that either $\omega = 0$ or $\omega = 1$. In view of (12), the relation (5) is equivalent to the pair of limit relations

$$r'/r \to 0 \text{ as } t \to \infty \tag{22}$$

and

$$r \to \infty \text{ as } t \to \infty \qquad \text{if } \omega = 0, \tag{23_0}$$

or

$$r^2 \to |c| \text{ as } t \to \infty \qquad \text{if } \omega = 1. \tag{23_1}$$

The relations (23) express the fact that

$$c^2/r^4(t) \to \omega \text{ as } t \to \infty. \tag{24}$$

If (10) is differentiated twice and (1) is used in the form $x'' = -fx$ and $y'' = -fy$, there results the relation $-r''/r = f - (xy' - yx')^2/r^4$. In view of (13), this means that $r = r(t)$ satisfies the "linear" differential equation

$$r'' + g(t)r = 0, \tag{25}$$

where

$$g(t) = f(t) - c^2/r^4(t). \tag{26}$$

The equation (25) is non-oscillatory, since it possesses the solution $r = r(t) > 0$. By Theorem B in [2], p. 391, a necessary and sufficient condition for (25) to possess a solution $r = r(t)$ satisfying (22) is that

$$\text{l. u. b.}_{0<s<\infty} \left| \int_t^{t+s} g(u)\,du \right| /(1+s) \to 0 \text{ as } t \to \infty. \tag{27}$$

Hence (15) follows from (24) and (26). This proves (I).

4. Proof of (II_0) and (III_0). The assertion (III_0) will be considered first. If (1) has a solution $z = z(t)$ of type (7), (13) satisfying (16_0), then (15_0) holds, by (I). In addition, the proof of (I) shows that (16_0) implies (23_0), that is, (17). Hence (15_0) and (17) are necessary for (16_0).

Conversely, if (15_0) and (17) are assumed, then (27) holds. Consequently, Theorem B of [2], p. 391, shows that the solution $r = r(t)$ of the non-oscillatory equation (25) satisfies (22). In view of (12), the relations (22) and (23_0) (that is, (17)) give (15_0). This proves (III_0).

In (II_0), it is assumed that (1) has a solution $z = z(t)$ of type (7), (13) satisfying (16_0). Thus, (15_0) holds, by (I). The assertion (III_0) implies, therefore, that a solution $z = z(t)$ of (7), (13) satisfies (16_0) if and

only if (17) holds. Consequently, it is sufficient to show that if a pair of real-valued, linearly independent solutions $x = x(t)$, $y = y(t)$ satisfies (17), then any pair of real-valued, linearly independent solutions, $x = c_1x + c_2y$ and $y = c_3x + c_4y$, where $c_1c_4 - c_2c_3 \neq 0$, satisfies (17), that is, $(c_1x + c_2y)^2 + (c_3x + c_4y)^2 \to \infty$ as $t \to \infty$. But this is clear, since $c_1c_4 - c_2c_3 \neq 0$ implies that there is a const. > 0 satisfying

$$(c_1x + c_2y)^2 + (c_3x + c_4y)^2 \geqq \text{const.}\ (x^2 + y^2)$$

for arbitrary x, y. This proves (II_0).

5. Proof of (II_1). Suppose that (1) has a pair of solutions $z_1 = z_1(t)$, $z_2 = z_2(t)$ satisfying $z_1'/z_1 \to i$, $z_2'/z_2 \to i$ as $t \to \infty$. Then

$$z_1'/z_1 - z_2'/z_2 = (z_1'z_2 - z_1z_2')/z_1z_2 \to 0 \text{ as } t \to \infty.$$

Here the Wronskian $z_1'z_2 - z_1z_2'$ is a constant. On the other hand, $|z_1|$ and $|z_2|$ tend to finite limits, as $t \to \infty$; cf. the derivation of (23_1) from (5). Thus, the relation in the last formula line cannot hold unless $z_1'z_2 - z_1z_2' \equiv 0$, that is, unless $z_1 = \text{const.}\, z_2$. This proves (II_1).

6. Proof of (III_1). The necessity of (15_1) and (18) for $(16_\pm)$ follows from the assertion (I) and its proof (cf. (23_1)).

In order to prove the sufficiency of (15_1) and (18), let $z = z(t)$ be a solution of (1) of type (7), (13) satisfying (18). The proof of (III_1) will be completed first under the assumption that

(28) $\qquad (15_1)$ and (18) imply (23_1).

It will then be shown that (28) is always true.

If (23_1) holds, then (27) follows from (15_1) and (26). In this case, (22) is a consequence of (25), (27); see [2], p. 391. Finally, $(16_\pm)$ follows from (12) and (22), (23_1).

Thus, in order to complete the proof of (III_1), only the assertion (28) remains to be proved. It should be noted that if (15_1) is replaced by (2) and (8), then (23_1) is a consequence of the fact that, on the one hand, the number $N(t)$ of zeros of $y(s)$ on $0 < s < t$ satisfies

(29) $\qquad \pi N(t) \sim t$

and, on the other hand,

(30) $\qquad \left| \pi N(t) - \int^t |c|\, ds/r^2(s) \right| \leqq \text{Const. as } t \to \infty;$

cf. [4], pp. 210-211.

In order to prove (28) (without the assumptions (2) and (8)), note that, since (25) is non-oscillatory,

$$\text{either } \int^t g(s)ds \to -\infty \text{ or } \lim_{t\to\infty} t^{-1}\int^t \left(\int^s g(u)du\right)ds \text{ exists}$$

(as a finite limit); cf. [2], p. 390. Hence (15_1) and (26) imply that

$$1 \leqq \lim_{t\to\infty} c^2/r^4(t) \ (\leqq\infty).$$

Suppose, if possible, that (23_1) does not hold, so that

$$1 < \lim_{t\to\infty} c^2/r^4(t) \ (\leqq\infty). \tag{31}$$

Then the solution $r = r(t)$ of (25) satisfies

$$1 + (r'/r)^2 \to \lim_{t\to\infty} c^2/r^4(t) \ (\leqq\infty) \text{ as } t\to\infty. \tag{32}$$

This follows from [1], pp. 570-571, if the limit in (31) is finite and from this result and from comparison theorems (cf. [5], pp. 635-636) if the limit is infinite. In view of (31), the relations (31), (32) imply that $r'/r < 0$ for large t; in fact, $r'/r < -\text{const.} < 0$ for large t. Hence

$$r < \exp(-\text{Const.}\, t) \text{ for large } t \qquad (\text{Const.} > 0). \tag{33}$$

By (13) and Schwarz's inequality, $c^2 \leqq r^2(t)(x'^2 + y'^2)$. Consequently,

$$x'^2 + y'^2 > c^2 \exp(2 \text{ Const.}\, t) \text{ for large } t. \tag{34}$$

If (1) is multiplied by x and if xx'' is written as $(xx')' - x'^2$, then a quadrature leads to

$$\int^t x'^2 ds = xx' + \int^t f(s)x^2(s)ds + \text{const.} \tag{35_1}$$

An integration by parts gives

$$\int^t fx^2 ds = x^2F - 2\int^t xx'F ds, \tag{35_2}$$

where $F(t) = \int^t f(s)ds$. Since $2\,|\,ab\,| \leqq \frac{1}{2}a^2 + 2b^2$ for any pair of real numbers a, b,

$$|\,2\int^t xx'F ds\,| \leqq \tfrac{1}{2}\int^t x'^2 ds + 2\int^t F^2x^2 ds. \tag{35_3}$$

The relations (35_1), (35_2), (35_3) give

$$\tfrac{1}{2}\int^t x'^2 ds \leqq xx' + O(t) + O(t^3), \text{ as } t \to \infty,$$

since $F(t) = O(t)$ and $|x| \leqq r(t) = O(1)$, by (15_1) and (31), respectively. Thus

$$\tfrac{1}{2}\int^t (x'^2 + y'^2) ds \leqq xx' + yy' + O(t^3) \text{ as } t \to \infty.$$

Another quadrature shows that

$$\tfrac{1}{2}\int^t \Big(\int^u (x'^2 + y'^2) ds\Big) du \leqq \tfrac{1}{2} r^2(t) + O(t^4) \text{ as } t \to \infty.$$

In view of (31), the term $r^2(t)$ is bounded as $t \to \infty$, and so the last relation contradicts (34). This contradiction shows that the hypothesis (31) is untenable. Hence (28) follows. This completes the proof of (III_1).

Remark. In this proof of (28), the full force of (18) was not used. Correspondingly, the proof remains applicable when $\lim |z(t)|$, as $t \to \infty$, exists as a finite limit or as ∞. In fact, whether or not $r(\infty)$ $(\leqq \infty)$ exists, the proof shows that

$$\liminf_{t\to\infty} c^2/r^4(t) \leqq 1 \leqq \limsup_{t\to\infty} c^2/r^4(t); \tag{36}$$

cf. (29), (30) for the case that (2), (8) hold.

7. Proof of (IV_1). Let $\rho = \rho(t)$ denote

$$\rho(t) = |z'(t)| = (x'^2 + y'^2)^{\frac{1}{2}} > 0. \tag{37}$$

Then a differentiation of (11) leads to $rr'' + r'^2 = \rho^2 + xx'' + yy''$. Hence, if (1) is written in the form $x'' = -fx$, $y'' = -fy$ and (25), (26) are used, it follows that

$$\rho^2 = r'^2 + c^2/r^2. \tag{38}$$

Since $(16_\pm)$ is equivalent to (22) and (23_1), it is seen that (19) is a necessary condition for $(16_\pm)$. Thus, all that remains to be shown is that (15_1) and (19) are sufficient for $(16_\pm)$. In view of (III_1), it is sufficient to show that

(39) (15_1) and (19) imply (18).

To this end, note that

$$\rho^2(\infty) = \lim_{t\to\infty} \rho^2(t) \geqq |c|. \tag{40}$$

In fact, it is clear from (38) and $r'^2 \geqq 0$ that $\limsup_{t\to\infty} c^2/r^2(t) \leqq \rho^2(\infty)$; so that (40) follows from the second inequality in (36).

If $\epsilon > 0$ and if t is sufficiently large, then $c^2/r^2(t) \leqq \rho^2(\infty) - 2\epsilon^2$ implies that $|r'(t)| > \epsilon$, by (38). If $t = t_0$ is sufficiently large and $c^2/r^2(t_0) \leqq \rho^2(\infty) - 2\epsilon^2$, then either $r'(t_0) > \epsilon$ or $r'(t_0) < -\epsilon$. In the former case,

$$c^2/r^2(t) \leqq \rho^2(\infty) - 2\epsilon^2,$$

hence $r'(t) > \epsilon$, for all $t \geqq t_0$. But $r'(t) > \epsilon$ for large t implies that $r(t) \to \infty$ as $t \to \infty$, and this contradicts (36). Consequently, $r'(t_0) < -\epsilon$ whenever $c^2/r^2(t_0) \leqq \rho^2(\infty) - 2\epsilon^2$ (and t_0 is sufficiently large). It follows that $c^2/r^2(t) > \rho^2(\infty) - 2\epsilon^2$ for all large t, that is, $\liminf c^2/r^2(t) \geqq \rho^2(\infty)$. Thus, by the last formula line, $\lim c^2/r^2(t)$ exists and is $\rho^2(\infty)$. Since $\rho^2(\infty) \geqq |c| > 0$, by (40), the relation (18) holds. This proves (39) and completes the proof of (IV_1).

8. Proof of (V_1). Let $z = z(t)$ be a solution of (1) of type (7), (13) satisfying (16_+). In view of (22) and (23_1),

$$rr' \to 0 \text{ and } r^2 \to c\ (>0) \text{ as } t \to \infty. \tag{41}$$

Consider the first of these two relations and the relation (13), that is,

$$xx' + yy' = o(1) \text{ and } yx' - xy' = -c,$$

as a pair of linear relations in x' and y'. Solving for x' and y', one obtains $(x^2+y^2)x' = -cy + o(|x|)$ and $(x^2+y^2)y' = cx + o(|y|)$, as $t \to \infty$. Hence, in view of the second relation in (41),

$$x' = -y + o(1) \text{ and } y' = x + o(1). \tag{42}$$

Thus (20) follows from (18). Similarly,

$$\lim_{t\to\infty} |y + iy'| \text{ exists.} \tag{20 bis}$$

Finally, an arbitrary real-valued solution, $c_1x + c_2y$, of (1) satisfies

$$\begin{aligned}(c_1x + c_2y)^2 + (c_1x' + c_2y')^2 \\ = c_1{}^2 |x + ix'|^2 + c_2 |y + iy'|^2 + 2c_1c_2(xy + x'y').\end{aligned} \tag{43}$$

In view of (42),

$$xy + x'y' = o(1) \text{ as } t \to \infty. \tag{44}$$

Hence

$$\lim_{t\to\infty} |(c_1x + c_2y) + i(c_1x' + c_2y')| \text{ exists,} \tag{45}$$

by (43), (20) and (20 bis). This proves that the condition that (20) holds for every real-valued solution $x(t)$ of (1) is necessary for the existence of a $z = z(t)$ satisfying (16_+).

Conversely, suppose that (20) holds for every real-valued solution $x = x(t)$ of (1). In particular, (45) holds for arbitrary constants c_1, c_2 and a pair of real-valued, linearly independent solutions, $x = x(t)$ and $y = y(t)$, of (1), with $c > 0$ in (13).

In view of (43), the assumption (45), for arbitrary constants c_1 and c_2, implies that $\lim (xy + x'y')$ exists. Without altering the value of c in (13), the solution $y(t)$ can be replaced by $bx + y$, where the constant b is arbitrary. Then $xy + x'y'$ becomes replaced by

$$x(bx + y) + x'(bx' + y') = b(x^2 + x_1'^2) + xy + x'y'.$$

Thus, if b is suitably chosen and $bx + y$ is renamed y, then (44) holds. If the pair of relations (13) and (44), which are linear in x, y', are solved for x, one obtains $(y^2 + y'^2)x = cy' + o(1)$. If the solutions x, y are replaced by const. x, y/const., then (13) and (44) remain unchanged and it can be supposed that $y^2 + y'^2 = c + o(1)$, $(c > 0)$, after a proper choice of const., if y/const. is renamed y. Thus $x = y' + o(1)$. But then (20 bis) gives (18), hence (16_-). In fact, (16_+) holds, since $c > 0$ in (13). This completes the proof of (V_1).

9. Examples to the case (8). There exist real-valued, continuous functions $f(t)$ which satisfy (2) and (8) but are such that (1) has no (complex-valued) solution $z = z(t)$ satisfying $(16_\pm)$. In fact, it was pointed out above that (9) is an example of such an $f(t)$.

As mentioned above, there exist continuous $f(t)$ satisfying (2) and (8), but such that (1) has no solution $z = z(t)$ satisfying $(16_\pm)$, although (21) holds for every solution $x = x(t)$ of (1). Such an $f(t)$ can be obtained by modifying the example in [5], pp. 631-632, as follows: Let $\lambda_1, \lambda_2, \cdots$ be a sequence of numbers satisfying

$$\lambda_k \to 1 \text{ as } k \to \infty, \tag{46}$$

put $\beta_1 = 0$ and $\beta_k = \frac{1}{2}\pi \sum_{j=1}^{k} 1/\lambda_j$ for $k > 1$, and define $f(t)$ by placing

$$f(t) = \lambda_k^2 \text{ if } \beta_{k-1} \leqq t < \beta_k \qquad (k = 1, 2, \cdots). \tag{47}$$

Thus $f(t)$ is defined for $0 \leqq t < \infty$. It is not continuous, but it will be clear that its discontinuities can be smoothed out (locally) without affecting the asymptotic behavior (as $t \to \infty$) of the solutions of (1).

Let $d_1, d_2, \cdots$ and $e_1, e_2, \cdots$ be the sequences of numbers

$$d_k = \prod_{j=1}^{k-1} \lambda_{2j}/\lambda_{2j+1} \text{ and } e_k = \prod_{j=1}^{k-1} \lambda_{2j+1}/\lambda_{2j+2}, \tag{48}$$

so that

$$e_k = \lambda_2/(d_k \lambda_{2k}). \tag{49}$$

On the half-line $\beta_2 \leqq t < \infty$, define a function $x = x_1(t)$ by the formulae

$$(-1)^k d_k \cos \lambda_{2k}(t - \beta_{2k}) \text{ or } (-1)^{k+1} d_{k+1} \sin \lambda_{2k+1}(t - \beta_{2k+1})$$

according as $\beta_{2k} \leqq t < \beta_{2k+1}$ or $\beta_{2k+1} \leqq t < \beta_{2k+2}$. On the half-line, $\beta_3 \leqq t < \infty$, define $x = x_2(t)$ by the formulae

$$(-1)^k e_k \cos \lambda_{2k+1}(t - \beta_{2k+1}) \text{ or } (-1)^{k+1} e_{k+1} \sin \lambda_{2k+2}(t - \beta_{2k+2})$$

according as $\beta_{2k+1} \leqq t < \beta_{2k+2}$ or $\beta_{2k+2} \leqq t < \beta_{2k+3}$. The functions $x_1(t)$, $x_2(t)$ have continuous first derivatives for $t \geqq \beta_2$, $t \geqq \beta_3$, respectively, and, in the case (47), satisfy (1) for all $t \neq \beta_1, \beta_2, \cdots$; so that x_1, x_2 are "solutions" of (1). It is clear that

$$|x_1(\beta_{2k})| = d_k, \ |x_1'(\beta_{2k})| = 0 \text{ and } |x_2(\beta_{2k+1})| = e_k, \ |x_2'(\beta_{2k+1})| = 0.$$

Hence, (20) cannot hold for $x = x_1$ and/or $x = x_2$, respectively, unless

$$\lim_{k \to \infty} d_k \text{ exists and/or } \lim_{k \to \infty} e_k \text{ exists.} \tag{50}$$

On the other hand, (21) holds for $x = x_1$ and $x = x_2$ (hence for every solution of (1)) if and only if

$$d_k = O(1) \text{ and } e_k = O(1) \text{ as } k \to \infty. \tag{51}$$

It is clear that the sequence $\lambda_1, \lambda_2, \cdots$ can be chosen so as to satisfy (46) and (51) but to violate (50). Since (46)-(47) implies (2) and (8), there results an $f(t)$ for which (1) has no solution $z = z(t)$ satisfying (16_+), although (21) holds for every solution.

It is also clear that the $f(t)$ in this example can be chosen so as to satisfy $f(t) \geqq 1$ for all t (or $f(t) \leqq 1$ for all t). On the other hand, *if* $f(t)$ *satisfies* (2), (8) *and if*

$$f(t) \text{ is monotone} \tag{52}$$

(*or just of bounded variation on* $0 \leqq t < \infty$), *then* (1) *has solutions* (7) *satisfying* (16_+). This follows from asymptotic formulae ([11], p. 270) for the solutions of (1).

10. Examples to the case (6). As mentioned above, Perron [7] has shown that there exist functions $f(t)$ which satisfy (2) and (6) but are such that (1) has no solution $z=z(t)$ satisfying (16_0). In view of (I_0) and (II_0), an $f(t)$ has these properties if it satisfies (2), (6) and if (1) is oscillatory and possesses a solution $x=x_0(t)\not\equiv 0$ satisfying

$$x_0(t)=o(1) \text{ as } t\to\infty \tag{53}$$

or merely

$$x_0(t)=O(1) \text{ as } t\to\infty. \tag{54}$$

For $|z(t)|=\text{const.}|x_0(t)|$ at the zeros of a fixed solution $x=y_0(t)$, linearly independent of $x_0(t)$ (while the zeros of $y_0(t)$ tend to ∞, since (1) is assumed to be oscillatory).

It follows that the example of [3], p. 534, has the desired properties and, in addition, (52) holds (the fact that $f(t)$ is a step-function, and not continuous, is unimportant for the considerations at hand). In other words, (2) *and* (52) *do not assure the truth of Poincaré's assertion concerning* (5) *in the case* (6), although, as mentioned above, (2) and (52) are sufficient for the truth of Poincaré's assertion in the case (8).

It is curious that in the example of Perron and in the example of [3], just mentioned, (1) has a solution satisfying (54), but not (53). In fact, if (2) and (6) hold and if (1) is oscillatory (so that f is positive and non-increasing), then no solution $x=x(t)$ $(\not\equiv 0)$ can satisfy (53) (since the relative maxima of $|x(t)|$ form a non-decreasing sequence); cf., e.g., [12].

It is therefore worth noting that $f(t)$ *can satisfy* (2), (6) *and have the property that* (1) *is oscillatory and has a non-trivial* $(\not\equiv 0)$ *solution satisfying* (53). An example of such an $f(t)$ is given in [5], pp. 643-644. The function given there even satisfies $f(t)=O(t^{-2})$ as $t\to\infty$ (hence $\int^{\infty}|f|dt<\infty$) and is such that (53) becomes improved to $x(t)=O(t^{\frac{1}{2}-\frac{1}{2}\gamma})$ as $t\to\infty$, where γ (>1) is a fixed (but arbitrary) number.

Note that (2), (6) prevent that (53) holds for *all* solutions of (1). In fact, the same is true if (53) is relaxed to (54). This follows from (30) and from the fact that (29) must be replaced by $N(t)=o(t)$ as $t\to\infty$ in the case (6); cf. [13].

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] P. Hartman, "Unrestricted solution fields of almost-separable differential equations," *Transactions of the American Mathematical Society*, vol. 63 (1948), pp. 560-580.

[2] ———, "On non-oscillatory linear differential equations of second order," *American Journal of Mathematics*, vol. 74 (1952), pp. 389-400.

[3] ——— and A. Wintner, "On non-conservative linear oscillators of low frequency," *ibid.*, vol. 70 (1948), pp. 529-539.

[4] ——— and A. Wintner, "A criterion for the non-degeneracy of the wave equation," *ibid.*, vol. 71 (1949), pp. 206-213.

[5] ——— and A. Wintner, "Oscillatory and non-oscillatory linear differential equations," *ibid.*, vol. 71 (1949), pp. 627-649.

[6] O. Perron, "Ueber lineare Differentialgleichungen, bei denen die unabhängige Variable reell ist," *Journal für die reine und angewandte Mathematik*, vol. 142 (1913), pp. 254-270.

[7] ———, "Ueber das infinitäre Verhalten der Integrale einer linearen Differentialgleichung zweiter Ordnung, wenn die characteristische Gleichung zwei gleiche Wurzeln hat," *Sitzungsberichte der Heidelberger Akademie*, 1917, Abhandlung 9.

[8] ———, "Ueber ein vermeintliches Stabilitätskriterium," *Göttinger Nachrichten*, 1930, pp. 28-29.

[9] H. Poincaré, *Œuvres*, vol. 1, 1928.

[10] A. Wintner, "The adiabatic linear oscillator," *American Journal of Mathematics*, vol. 68 (1946), pp. 385-397.

[11] ———, "Asymptotic integrations of the adiabatic oscillator," *ibid.*, vol. 69 (1947), pp. 251-272.

[12] ———, "Stability and high frequency," *Journal of Applied Physics*, vol. 18 (1947), pp. 941-942.

[13] ———, "A criterion of oscillatory stability," *Quarterly of Applied Mathematics*, vol. 7 (1949), pp. 115-117.

ISOMORPHISMS BETWEEN PROJECTIVE UNITARY GROUPS.*[1]

By JOHN H. WALTER.

I. Introduction and Preliminaries.

1.1. Introduction. The problem of finding the automorphisms of the classical groups and the related problem of determining isomorphisms between particular groups has been studied by Dieudonné [5] and Rickart [7, 8] among others. With regard to the projective unitary groups, this question has been unanswered except when the underlying vector spaces are over finite fields (Dieudonné [5, p. 82]; the proof given there for the proper projective unitary groups applies also to the projective unitary groups). We shall treat this problem here.

Denote by E a right vector space of dimension n over a sfield K of characteristic not 2. Assume furthermore that K possesses an involutive anti-automorphism $J: \phi \to \phi^J$. Then we can define a sesquilinear form f as a mapping $f: E \times E \to K$ such that $f(x, y)$ is additive in both x and y and $f(x\alpha, y\beta) = \alpha^J f(x, y)\beta$ where α and β are in K. The form f is said to be hermitian if $f(y, x)^J = f(x, y)$ and nondegenerate if $f(x, y) = 0$ for all y in E implies that $x = 0$. We shall assume henceforth that f is a non-degenerate herimtian sesquilinear form.[2]

The group $U_n(K, f)$ of the linear transformations of E onto itself which leave invariant the form f is called the *unitary group*. The inner automorphism $\phi \to \lambda^{-1}\phi\lambda$ of K induced by the element λ will be denoted by λ^*; so $\phi^{\lambda^*} = \lambda^{-1}\phi\lambda$. The transformation $h_\lambda: x \to x\lambda$ for all x in E is semilinear with respect to the automorphism λ^* of K and is called a *homothetic transformation*. The group of homothetic transformations is denoted by $H_n(K)$.

* Received June 8, 1955.

[1] This paper is based on the author's dissertation submitted for the degree of Doctor of Philosophy at the University of Michigan. The author acknowledges the guidance and encouragement given by Professor J. A. Dieudonné. In part, this work has been sponsored by the Office of Ordnance Research, U. S. Army, under Contract DA-20-018-ORD 13281.

[2] Birkhoff and von Neumann [1] have shown that if f is reflexive and not alternating, then there exists an element α in K such that αf is hermitian. The projective group associated with these forms then will be identical to the projective unitary groups.

The factor group $U_n(K,f)/H_n(K) \cap U_n(K,f)$ is the *projective unitary group* $PU_n(K,f)$.

Dieudonné [5] and Rickart [9] have shown that if certain involutions of $PU_n(K,f)$, called extremal involutions, which are defined by the way they act on E and which are associated with the 1-dimensional nonisotropic subspaces of E can be characterized strictly in terms of the multiplicative structure of $PU_n(K,f)$, then the automorphisms of $PU_n(K,f)$ can be determined and the related question of determining the isomorphisms between certain subgroups of $PU_n(K,f)$ and similar subgroups of another projective unitary group can be answered. The group-theoretical characterization of extremal involutions is our principal result.

Besides the aforesaid restriction on K, we assume, whenever we use Lemma 1.5, that K is not the finite field $GF(3)$, a case which has been considered by Dieudonné [5]. For the sake of convenience, we furthermore assume that n is finite although the extension to the infinite dimensional case is immediate due to some remarks of Rickart [8; p. 700]. We shall ultimately restrict n to be greater than 4 but we do not assume this from the outset. Of course, we take $n \geq 2$ in all cases.

1.2. Definitions and Notations.[3] In the study of the projective unitary group, we shall consider also the group of collineations $\Gamma L_n(K)$ which consists of all 1-1 semilinear transformations of E, and the factor group $P\Gamma L_n(K) = \Gamma L_n(K)/H_n(K)$, which is called the group of projective collineations. By the group $\Gamma U_n(K,f)$ of semisimilitudes, we understand the subgroup of $\Gamma L_n(K)$ consisting of those elements t which satisfy the equation

$$f(t(x),t(y)) = e_t f(x,y)^\tau$$

where τ is the automorphism of K associated with t and e_t is an element of K, which depends only on t and is called "the multiplier of t." It can be shown that if ϕ is in K, $\phi^{\tau J} = e_t \phi^{J\tau} e_t^{-1}$. A homothetic transformation h_λ is contained in $\Gamma U_n(K,f)$ with multiplier $\lambda^J\lambda$. The factor group

$$P\Gamma U_n(K,f) = \Gamma U_n(K,f)/H_n(K)$$

is called the group of projective semisimilitudes.

If K is commutative and J is the identity, the groups $U_n(K,f)$ and $PU_n(K,f)$ become orthogonal and projective orthogonal groups $O_n(K,f)$

[3] The reader is referred to Dieudonné [2] from which this paper takes its point of departure. In places where details have been omitted, the reader is referred to Dieudonné's treatise.

and $PO_n(K,f)$, respectively, which we merely consider as special cases of the unitary and projective unitary groups.

The symplectic groups will be introduced. Such a group is defined, when K is commutative, as the group of linear transformations leaving invariant an alternating bilinear form f and will be denoted by $Sp_n(K,f)$. The corresponding group of semilinear transformations, however, will still be denoted by $\Gamma U_n(K,f)$.

With the exception of the case where $U_n(K,f)$ is the orthogonal group $O_2(K,f)$, K is the finite field $GF(3)$, and f is of index 1, the center of $U_n(K,f)$ is $C_n = H_n(K) \cap U_n(K,f)$. Let Z denote the center of K; then C_n consists of those homothetic transformations h_λ, where λ is in and $\lambda^J\lambda = 1$. The set of such elements λ forms a multiplicative group of Z which we denote by C. At times, a homothetic transformation h_λ will be denoted by λ and $h_\lambda(x)$ by $\lambda(x)$.

1.3. Classification of the involutions of $PU_n(K,f)$.

We shall use the methods of linear algebra and for that reason we look upon an element t^* of $PU_n(K,f)$ as a coset in $U_n(K,f)$. In particular, if u^* is an involution[4] of $PU_n(K,f)$, then an element u of u^* will be called a *projective involution*. If u^* is an involution of $P\Gamma U_n(K,f)$ then u^* will be called a *semilinear involution* and u of $\Gamma U_n(K,f)$ in the coset u^* will be called a *projective semilinear involution*. The relation $u^2 = \gamma$ with γ in C_n characterizes projective involutions.

We shall now distinguish various types of projective involutions u, where $u^2 = \gamma$ in C_n:

(i) u is of the *first kind* if $\gamma = \lambda^2$ when λ is in Z and $\lambda^J\lambda = 1$.

(ii) u is of the *second kind* if $\gamma = \lambda^2$ when λ is in Z and $\lambda^J\lambda = -1$.

(iii) u is of the *third kind* if $\gamma = \lambda^2$ when λ is not in Z.

(iv) u is of the *fourth kind* if γ is not a square in K.

Note that if $\gamma = \lambda^2$ and λ is in Z, then $\gamma^J\gamma = 1$ implies that $\lambda^J\lambda = \pm 1$. So in this case u is either of the first or second kind. If a projective involution u is of the k-th kind, then so is every projective involution ηu, where η is in C_n. So we can say that the involution u^* of $PU_n(K,f)$ which contains u is also of the k-th kind. An extremal involution is a particular kind of involution of the first kind. Because the identity 1 can obviously be dis-

[4] By an involution, we mean an element u^* such that $u^{*2} = 1$.

tinguished group-theoretically from other involutions, we shall exclude it from the set of involutions of the first kind and the elements of C_n from the set of projective involutions of the first kind.

It is clear that if two involutions in $U_n(K,f)$ or two involutions in $PU_n(K,f)$ are conjugate, they are of the same kind.

1.4. Definition of the invariants ν_u^* and ω_u^*. We shall use in our characterization of extremal involutions of $PU_n(K,f)$ group-theoretical invariants similar to those introduced by Mackey [6] and effectively used by Rickart [7, 8]. Let $\mathcal{S}^*$ be a set of involutions in $PU_n(K,f)$ and define $c(\mathcal{S}^*)$ to be the set of those involutions of $PU_n(K,f)$ which commute with every element of $\mathcal{S}^*$. Denote by

$\nu(u^*, v^*)$ the number of elements in $c(c(u^*, v^*))$;

ν_{u^*} the maximum of $\nu(u^*, v^*)$ for all involutions $v^* \neq u^*$ such that $u^*v^* = v^*u^*$, $u^* \neq 1^*$, and $v^* \neq 1^*$.

It is easily seen that $c(c(u^*, v^*))$ always contains 1^*, u^*, v^*, and u^*v^* when u^* and v^* are chosen as above; so $\nu_{u^*} \geqq 4$. Rickart [7, 8] showed in the cases he considered that the minimum of the invariant corresponding to ν_{u^*} was attained only for extremal involutions. In the case of the projective unitary groups, this is no longer true. So we introduce a new invariant.

Let $\mathfrak{A}^*$ denote the set of involutions of $PU_n(K,f)$ for which $\nu_{u^*} = 4$. As ν_{1^*} is not defined, 1^* is not in $\mathfrak{A}^*$. With one exception when $n = 6$, we shall show that $\mathfrak{A}^*$ contains the extremal involutions and the involutions of the third and fourth kinds. Denote by

$\omega(u^*, v^*, w^*)$ the number of elements of $c(c(u^*, v^*, w^*))$;

ω_{u^*} the maximum of $\omega(u^*, v^*, w^*)$ subject to the conditions that u^*, v^*, and w^* are distinct mutually commuting involutions in $\mathfrak{A}^*$ such that $u^*v^* \neq w^*$.

The condition that $u^*v^* \neq w^*$ implies that $\omega_{u^*} \geqq 8$. We have been able to show that the invariant ω_{u^*} alone distinguishes extremal involutions when $n > 4$ and $n \neq 6$, 8, or 12. However, in order to include the cases $n = 6$, 8 and 12, we do not rely completely on this invariant in this paper.

If u^* and v^* are commuting involutions, Dieudonné [5, p. 8] has shown that for u in u^* and v in v^*, $uv = \pm vu$ and that the choice of sign is independent of the choice of representatives. It is convenient to denote the set of elements of $U_n(K,f)$ determined by a set $\mathcal{S}^*$ of $PU_n(K,f)$ by $\mathcal{S}$. Conversely, if $\mathcal{S}$ is a set of elements in $U_n(K,f)$, we write $\mathcal{S}^*$ for the set of

elements in $PU_n(K,f)$ which have, as cosets, representatives in $\mathcal{S}$. Likewise, when $\mathcal{S}$ is a set of projective involutions, let $c(\mathcal{S})$ and $c(c(\mathcal{S}))$ denote the sets of projective involutions determined by $c(\mathcal{S}^*)$ and $c(c(\mathcal{S}^*))$, respectively. Then $c(\mathcal{S})$ can be characterized as the set of projective involutions which commute or anticommute with the elements of $\mathcal{S}$.

1.5. Some concepts in linear algebra. The usual concepts of orthogonality, isotropic, and nonisotropic subspaces, etc., will be used. By $S^\perp$, we mean the subspace orthogonal to a given subspace S. A 2-dimensional subspace will sometimes be called a plane.

We shall say that a 1-dimensional subspace X of a vector space E is *f-commensurable* (or in short, *commensurable*) to a 1-dimensional space Y if there exists a linear t of E mapping X onto Y such that $f(t(x),t(x))=f(x,x)$ for all x in X. We shall say that the vector x is *f-commensurable* to the vector $t(x)$. The relation of commensurability is obviously an equivalence relation. A subspace U of E is said to be *f-commensurable* to a subspace V if $\dim U=\dim V$ and the 1-dimensional nonisotropic subspaces of U are commensurable to 1-dimensional subspaces of V.

LEMMA 1.1. *Suppose that E can be decomposed into the direct sums*

$$E=U_1\oplus U_2\oplus\cdots\oplus U_k \text{ and } E=V_1\oplus V_2\oplus\cdots\oplus V_k$$

of orthogonal subspaces such that U_i is commensurable to V_i, $i=1,2,\cdots,k$. Then there exists an element t in $U_n(K,f)$ such that $u(U_i)=V_i$, $i=1,2,\cdots,k$.

LEMMA 1.2. *Suppose that E can be decomposed into the direct sum $E=U_1\oplus\cdots\oplus U_h\oplus V_1\oplus\cdots\oplus V_h\oplus W$ of orthogonal subspaces such that U_i is commensurable to V_i, $i=1,2,\cdots,h$. Then there exists an involution u in $U_n(K,f)$ such that $u(U_i)=V_i$, $i=1,2,\cdots,h$ and u leaves fixed W.*

The proofs of these lemmas are obvious and will be left to the reader.

We now say that E is a *f-commensurable* (or, in short, *commensurable*) vector space if all its 1-dimensional nonisotropic subspaces are commensurable. It follows that any two nonisotropic subspaces of the same dimension which are contained in a commensurable vector space are commensurable to each other.

LEMMA 1.3. *Let $E=U\oplus V$ where U and V are orthogonal nonisotropic subspaces which are commensurable to each other. If $\dim U=\dim V\geqq 3$, E is commensurable.*

Let z be a nonisotropic vector of E not in U or V. We shall prove the lemma by showing that zK is commensurable to a nonisotropic subspace of U or V. Indeed, $z = x_1 + x_2$, where x_1 is in U and x_2 is in V. Not both x_1 and x_2 can be isotropic; for then z would be isotropic. Suppose x_1 is nonisotropic. If x_2 is nonisotropic let $Q = x_2K$ and pick y in $Q^\perp \cap V$ so that y is nonisotropic. If x_2 is isotropic, let Q be a nonisotropic plane contained in V and containing x_2. Then, $\dim Q^\perp \cap V \geqq 1$; so again pick y in $Q^\perp \cap V$ so that y is nonisotropic.

Let $X = x_1K$, $X' = X^\perp \cap U$, $Y = yK$, $Y' = Y^\perp \cap V$. Then

$$E = X \oplus X' \oplus Y \oplus Y'.$$

Hence by Lemma 1.2, there exists an involution u of $U_n(K,f)$ such that $u(x_1) = y\alpha$, where α is in K, and u leaves fixed X' and Y'. But then $u(z) = y_1\alpha + x_2$ is in V; so zK is commensurable to $(y_1\alpha + x_2)K$ in V.

1.6. Involutions of the first kind.[5] Let u^* be an involution of the first kind in $PU_n(K,f)$ and choose a representative u in u^*. Then $u^2 = \gamma$ and $\gamma = \lambda^2$, where λ is in C. The projective involution $\lambda^{-1}u$ is easily seen to be an involution in u^*; so we may consider u to be an involution. Then E is the direct sum $U^+ \oplus U^-$ of the eigenspaces of u, where $u(x) = \pm x$ according as x is in U^+ or U^-, respectively. The subspaces U^+ and U^- are orthogonal and nonisotropic (Dieudonné [4; p. 19]). If $\dim U^+ = p$, u is said to be a *p-involution*. As both u and $-u$ are in u^*, we can always choose u so that $p \leqq [n/2]$, where $[r]$ designates the greatest integer less than or equal to a rational number r. With this choice of representative, we also call u^* a *p-involution*. Those linear involutions for which $p = 1$ or $p = n - 1$ are called *minimal involutions* and the corresponding cosets are called *extremal involutions*.

An element t in $\Gamma L_n(K)$ commutes with an involution u if and only if t leaves invariant the eigenspaces of u. If u and v are commuting involutions of $U_n(K,f)$ with eigenspaces $U^\pm$ and $V^\pm$, respectively, then $E = P_1 \oplus P_2 \oplus P_3 \oplus P_4$, where $P_1 = U^+ \cap V^+$, $P_2 = U^+ \cap V^-$, $P_3 = U^- \cap V^+$, and $P_4 = U^- \cap V^-$. The subspaces P_i are orthogonal and nonisotropic when they are not zero. If t in $\Gamma L_n(K)$ anticommutes with u, then t maps the eigenspaces U^+ and U^- onto one another. In this case, $\dim U^+ = \dim U^- = n/2 = m$ That is, *no element t of $\Gamma L_n(K)$ can anticommute with an involution u of $U_n,K,f)$ except if u is a m-involution.*

---

[5] Dieudonné [2] has characterized the centralizers of projective involutions in $\Gamma U_n(K,f)$. The next four sections are an adaptation of this characterization to our case.

If two 1-involutions commute, their 1-dimensional eigenspaces are orthogonal. Thus a maximal set of commuting 1-involutions consists of n involutions whose 1-dimensional eigenspaces are mutually orthogonal. When $n > 2$, no element of $U_n(K,f)$ anticommutes with a 1-involution. Consequently a maximal set of commuting 1-involutions of $U_n(K,f)$ determines a maximal set of commuting extremal involutions of $PU_n(K,f)$ which always contains n involutions. The product of p mutually commuting 1-involutions of $U_n(K,f)$ is easily seen to be a p-involution. Therefore, the product of p mutually extremal involutions is a p-involution or $(n-p)$-involution according as $p \leqq [n/2]$ or $n-p \leqq [n/2]$.

When E is commensurable, it is a direct consequence of Lemma 1.1 that, for each p, all p-involutions in $U_n(K,f)$ are conjugate. Then, for each p, all p-involutions of $PU_n(K,f)$ are conjugate.

1.7. Involutions of the second kind. Let u^* be an involution of the second kind in $PU_n(K,f)$ and choose a representative u in u^*. Then $u^2 = \gamma$, where $\gamma = \lambda^2$, λ is in Z, and $\lambda^J\lambda = -1$. Then $u' = \lambda^{-1}u$ is a linear involution of $\Gamma U_n(K,f)$ with multiplier -1 and $E = U^+ \oplus U^-$, where U^+ and U^- are the eigenspaces of u' and of U for any u in u^*. The subspaces U^+ and U^- are totally isotropic of dimension $m = n/2$; consequently n must be even [3, p. 225]. The conditions for t in $\Gamma L_n(K)$ to commute with u are the same as for involutions of the first kind. If v is a projective involution of the first or second kind, then we have $E = P_1 \oplus P_2 \oplus P_3 \oplus P_4$, where the subspaces P_i are defined in Section 1.6.

1.8. Involutions of the third kind. We shall consider the following generalization of projective involutions of the third kind, which arises in the subsequent developments. Let u be a projective semilinear involution of $\Gamma U_n(K,f)$ relative to an involutive automorphism σ with multiplier e such that $u^2 = \gamma$ in C_n, where $\gamma = \lambda\lambda^\sigma$ for some λ in K. In case $\lambda^\sigma = \lambda$, we furthermore assume that λ is not in Z. Then $u' = \lambda^{-1}u$ is a semilinear involution of $\Gamma U_n(K,f)$ relative to the automorphism $\sigma' = (\lambda^{-1})^*\sigma = \sigma\lambda^*$ and with multiplier $e' = (\lambda\lambda^J)^{-1}e$. It follows that $\sigma'^2 = 1$ and $\sigma' \neq 1$. Let K_1 be the fixed subsfield of K by σ'. Then there exists an element θ in K such that $\theta^{\sigma'} = -\theta$ and K is the direct sum of the two K_1-spaces: $K = K_1 \oplus K_1\theta$. Thus E can be considered as a right vector space E^* over K_1 and $\dim_{K_1} E^* = 2n$. As a transformation of E^*, u' is a linear involution. So $E^* = U^+ \oplus U^-$, where U^+ and U^- are the eigenspaces of u'. Furthermore $U^-\theta = U^+$; so $\dim_{K_1} U^+ = \dim_{K_1} U^- = n$ and $E^* = U^+ \oplus U^+\theta$. Thus a K_1-basis for U^+ is also a K-basis for E. We shall sometimes denote U^+ as E_1.

Now we associate with f another nondegenerate sesquilinear form defined on E_1. In general, we have $\phi^{J\sigma'} = e'^{-1}\phi^J e'$ for ϕ in K_1. Dieudonné [2; chap. I, § 14] shows that e can always be taken to be 1 by multiplying the form f by an element d^{-1} of K such that $e' = d^{1-\sigma}$ and $d^J = \pm d$. In fact, if $e' \neq -1$, it suffices to take $d = 1 + e'$. If $e' = -1$ and there exists an element ψ of $K_1\theta$ such that $\psi^J \neq -\psi$, take $d = \psi + \psi^J$. In both these cases $d^J = d$ and so the form $g = d^{-1}f$ is a hermitian sesquilinear form relative to the anti-automorphism Jd^*.

However, if $\psi^J = -\psi$ for all elements ψ of $K_1\theta$ and if ϕ is any element of K_1, we have $(\phi\psi)^J = -\phi\psi = \psi^J\phi^J = -\psi\phi^J$. So the restriction of J to K_1 is the automorphism $\phi \to \psi^{-1}\phi\psi$ where ψ is any element of $K_1\theta$; hence K_1 is commutative.[6] Now take any element of $K_1\theta$ for d. Then $d^{1-\sigma'} = -1 = e$. But now g is a skew hermitian sesquilinear form relative to the antiautomorphism $T = Jd^*$. Furthermore, since d^2 is in K_1, $\phi^T = d^{-1}\phi^J d = d^{-2}\phi d^2 = \phi$ for ϕ in K_1 and thus the restriction of T to K_1 is the identity.

Now we have that $\Gamma U_n(K, g) = \Gamma U_n(K, f)$ and $U_n(K, g) = U_n(K, f)$. In $\Gamma U_n(K, g)$, u' has multiplier 1; so if x and y are in E_1,

$$g(x, y) = g(u'(x), u'(y)) = g(x, y)^{\sigma'}$$

and $g(x, y)$ is in K_1. Thus we can restrict g to a sesquilinear form f_1 such that $f_1 : E_1 \times E_1 \to K_1$. In case g is hermitian, f_1 is hermitian and we form the groups $U_n(K_1, f_1)$ and $\Gamma U_n(K_1, f_1)$ acting on E_1.

In case g is skew-hermitian, f_1 is alternating and we form the group $Sp_n(K_1, f_1)$. Let x and y be in E_1 such that $g(x, y) = f_1(x, y) = 1$. Then $g(x, y) = g(u(x), u(y)) = \lambda^T\lambda g(x, y)^{\lambda^*}$. So $\lambda^T\lambda = 1$. In case $\lambda^\sigma = \lambda$, λ is in K_1 and we have $\lambda^2 = 1$. But we have excluded the case that $\lambda^\sigma = \lambda$ and λ is in Z. Hence f_1 can be alternating only if $\lambda^\sigma \neq \lambda$; and thus when u is linear, this case is excluded.

An element t of $\Gamma U_n(K, f)$ relative to an automorphism τ, where $\lambda^\tau = \lambda$ and with multiplier e_t which commutes with u commutes also with u'. It then follows that $\sigma'\tau = \tau\sigma'$; so $K_1{}^\tau = K_1$. Thus t can be restricted to a semi-linear transformation t_1 acting on E_1. Furthermore, t is in $\Gamma U_n(K, g)$ with multiplier $e_t' = d^{-1}e_t d$. But then

$$g(u't(x), u't(y)) = e_t'^{\sigma'}g(x, y)^{\sigma'\tau} = g(tu'(x), tu'(y)) = e_t g(x, y)^{\sigma'\tau}.$$

Since $\tau\sigma' = \sigma'\tau$, $e_t'^{\sigma'} = e_t'$ and e_t' is in K_1. Thus the restriction t_1 of t to E_1 is in $\Gamma U_n(K_1, f_1)$ or $SP_n(K_1, f_1)$ relative to the restriction of τ to K_1, as the case may be. We shall say that t_1 *represents* t in $\Gamma U_n(K_1, f_1)$.

---

[6] K can be shown to be either commutative or a sfield of generalized quaternions in this case.

In general, an element of $\Gamma U_n(K_1, f_1)$ cannot be extended to an element t of $\Gamma U_n(K, f)$. But if t_1 is in $U_n(K_1, f_1)$ or $Sp_n(K_1, f_1)$ this can be done by setting $t(y) = t_1(x)\theta$ for y in U^-, where x is the vector in $U^+ = E_1$ such that $y = x\theta$.

To summarize we have the following remark.

An element t of $\Gamma U_n(K, f)$ which is relative to an automorphism τ such that $\lambda^\tau = \lambda$ and which commutes with the projective semilinear involution u is represented by its restriction t_1 to E_1 in $\Gamma U_n(K_1, f_1)$. An element of $U_n(K_1, f_1)$ or $Sp_n(K_1, f_1)$ represents an element of $U_n(K, f)$ which commutes with u.

Let Z_1 be the center of K_1. For ϕ in K_1, we have $\phi^{\theta^*\sigma'} = \theta^{-\sigma'}\phi^{\sigma'}\theta^{\sigma'} = \phi^{\theta^*}$; so θ^* leaves K_1 and Z_1 invariant. Since θ^2 is in K_1, the restriction of θ^* to Z_1 is an involutive automorphism of Z_1. Thus either every element of Z_1 commutes with θ (in case the restriction of θ^* is the identity) or by Galois theory $Z_1 = Z_0 \oplus Z_0\lambda'$, where $\lambda'\theta = -\theta\lambda'$ and every element of Z_0 commutes with θ.

Let Z' be the subfield of Z_1 which is left fixed by θ^*. Since $K = K_1 \oplus K_1\theta$, $Z' \subseteq Z$. On the other hand, $K_1 \cap Z \subseteq Z'$. So as $Z' \subseteq K_1$, $Z' = K_1 \cap Z$. Therefore if θ^* is the identity on Z_1, $Z_1 = K_1 \cap Z$. If θ^* is order 2 on Z_1 with fixed subfield Z_0, $Z_0 = K_1 \cap Z$.

In the important case where u is a projective involution of the third kind in $U_n(K, f)$, $\sigma' = (\lambda^{-1})^*$ and so λ is in Z_1. But $\lambda^{\theta^*} = -\lambda$. So θ^* is not the identity on Z_1. We can take $\lambda' = \lambda$. Then $Z_1 = Z \oplus Z\lambda$ as $K_1 \cap Z = Z$; that is, $Z_1 = Z(\lambda)$.

In general, when σ' leaves Z_1 fixed and θ^* does not, we have $Z_1 = Z(\lambda')$. In this case, K_1 is an interior Galois subsfield of K, and, according to a result by Dieudonné [3, p. 182], $K = K_1 \oplus K_1\psi$, where ψ is an element of K whose square $\psi^2 = \alpha$ is in Z. However, this implies that $\psi^{\sigma'} = -\psi$. Indeed $\psi^{\sigma'} = a + b\psi$ with a and b in K_1. Therefore $\psi = \psi^{\sigma'^2} = a(b + 1) + b^2\psi$. Thus $b = \pm 1$. If $b = 1$, then $a = 0$ and $\psi^{\sigma'} = \psi$, which contradicts the fact that ψ is not in K_1. Therefore, $\psi^{\sigma'} = a - \psi$ and

$$(\psi^{\sigma'})^2 = a^2 + \psi^2 - (\psi a + a\psi) = (\psi^2)^{\sigma'} = \alpha.$$

This implies $a = 0$ and $\psi^{\sigma'} = -\psi$. Since θ was chosen arbitrarily to satisfy the condition $\theta^{\sigma'} = -\theta$, we can henceforth assume that $\theta = \psi$ so that $\theta^2 = \alpha$ is in Z and $(\theta^*)^2 = 1$.

Now we consider the case that t in $\Gamma U_n(K, f)$ anticommutes with u, where u is now a projective involution of the third kind in $U_n(K, f)$. Let τ

be the automorphism of K associated with t and let e_t be the multiplier of t. Assume $\lambda^\tau = \lambda$. Then t commutes with $h_{\lambda^{-1}}$ and hence anticommutes with u'. As $\theta^{\sigma'} = -\theta$, $h_\theta u' = -u' h_\theta$. So $t' = h_\theta t$ is a projective semilinear involution of $\Gamma U_n(K,f)$ relative to the automorphism $\tau\theta^*$ and $t'u' = u't'$. The multiplier of t' in $\Gamma U_n(K,g)$ is $d^{-1}\theta^J e_t \theta d^{\tau\theta^*}$ which again must be in K_1 by the argument given previously in this section. Furthermore $K_1^{\tau\theta^*} = K_1$; so t' can be restricted to a transformation t_1 of $\Gamma U_n(K_1, f_1)$ acting on E_1. We say that in this case t_1 *represents* t.

In summary: *Let t be a projective semilinear involution of $\Gamma U_n(K,f)$ relative to an automorphism τ of K. Then if $ut = -tu$ and $\lambda^\tau = \lambda$, t is represented by a transformation t_1 in $\Gamma U_n(K_1, f_1)$, which is semilinear with respect to the automorphism $\tau\theta^*$.*

1. 9. Involutions of the fourth kind. Let u^* be an involution of the fourth kind in $PU_n(K,f)$ and choose a representative u in u^*. Then $u^2 = \gamma$ in C_n, where now γ is not a square in K. Again we consider a generalization. Let u be a projective semilinear involution in $\Gamma U_n(K,f)$ relative to an involutive automorphism σ of K and with multiplier e such that $u^2 = \gamma$, where γ is in C and is not of the form $\lambda\lambda^\sigma$ for λ in K.

Dieudonné [2, chap. I, § 13] has shown that there exists a quadratic extension L of K obtained by adjoining a square root ρ of γ such that $\phi\rho = \rho\phi^\sigma$ for ϕ in K. (Obviously L is not commutative when $\sigma \neq 1$, even though K may be. So the consideration of sfields is necessary at this point.) Now E can be considered as a right vector space F over L by setting $x\rho = u(x)$ for x in E. The L-dimension of F is $m = n/2$. Consequently n must be even, if this type of projective involution is to exist in $\Gamma U_n(K,f)$. The antiautomorphism J can be extended to L by setting $\rho^J = \rho e^\sigma \gamma^{-1} = e\rho^{-1}$. Corresponding to f, we define a new nondegenerate hermitian sesquilinear form g by setting $g(x,y) = f(x,y) + \rho f(x, u(y))^\sigma \gamma^{-1}$. In $\Gamma U_n(L,g)$ u is the homothetic transformation h_ρ with multiplier $e = \rho^J\rho$.

It is easily seen that if t is an element of $\Gamma U_n(K,f)$ which is relative to an automorphism τ of K and which commutes with u, then t can be considered to be in $\Gamma U_m(L,g)$ relative to the extension of the automorphism τ to L obtained by setting $\rho^\tau = \rho$. Conversely, if t is an element of $\Gamma U_n(L,g)$ which is relative to an automorphism τ such that $K^\tau = K$, then t can be considered as an element of $\Gamma U_n(K,f)$ which is relative to the restriction of τ to K and which commutes with u.

Let σ_0 be the automorphism of L leaving K fixed which is defined by the relation $\sigma_0 : \rho \to -\rho$. Clearly σ_0 commutes with any automorphism τ

of L such that $K^\tau = K$ and $\rho^\tau = \eta\rho$ with η in K. Therefore, if t is an element of $\Gamma U_n(K,f)$ which is relative to an automorphism τ of K and which anti-commutes with u, t can be considered to be in $\Gamma U_m(L,g)$ relative to the automorphism $\sigma_0\tau = \tau\sigma_0$, where now τ is extended to L by setting $\rho^\tau = \rho$. Conversely, any element t of $\Gamma U_m(L,g)$ which is relative to an automorphism of the form $\sigma_0\tau$, where $\rho^\tau = \rho$ and $K^\tau = K$, can be considered as an element of $\Gamma U_n(K,f)$ which is relative to the restriction of τ to K and which anti-commutes with u.

Denote the center of L by Z_L. If $\phi = \xi + \rho\eta$ is in Z_L with ξ and η in K, the condition that ϕ should commute with every element of K implies that ξ and η are in Z. Hence $Z \subseteq Z_L \subseteq Z(\rho)$. Now if $\sigma = 1$, ρ is in $Z_L = Z(\rho)$. However, if $\sigma \neq 1$, $Z_L = Z$ as $Z(\rho)$ is a quadratic extension of Z.

In Section 4.2, we make reference to the case where $U_n(K,f)$ is the orthogonal group $O_n(K,f)$ and u is a projective involution of the fourth kind in $O_n(K,f)$. Then K is commutative and J is the identity by definition of $O_n(K,f)$. If $\gamma = u^2$, we have now $\gamma^J\gamma = \gamma^2 = 1$. As γ is not a square in K, $\gamma = -1$, $\rho = i$ and $L = K(i)$, where $i = \sqrt{-1}$. In this case, the extension of the identity automorphism J to L is obtained by setting $i^J = i^{-1} = -i$.

1.10. A result on fields. The following remark is a special case of a well-known result on fields.

Let Z' be the splitting field of an irreducible equation $x^2 - \alpha = 0$ over a field Z. Then if β is an element of Z' such that β^2 is in Z, either β is in Z or $\beta = \lambda\rho$, where ρ is a root of $x^2 - \alpha = 0$ and λ is in Z.

1.11. Some preliminary lemmas. The lemmas of this section apply to sfields K which will be identified with the sfield, subsfields, or extension sfields which we associate with the projective unitary group under investigation in the following sections.

LEMMA 1.4. *If K is not of characteristic 2, then for every pair of orthogonal nonisotropic vectors x and y and for any α in K, at least one of the vectors $x + y(\alpha - 1)$, $x + y\alpha$, and $x + y(\alpha + 1)$ is nonisotropic.*

Clearly, as x and y are nonisotropic, f is not alternating and hence is hermitian. Compute first

$$f(x + y\beta, x + y\beta) = f(x,x) + \beta^J f(y,y)\beta. \tag{1.1}$$

The lemma is trivial if $\alpha = 0$ or $\alpha = \pm 1$. So suppose $\alpha \neq 0$ and $\alpha \neq \pm 1$

and that $x+y\alpha$ and $x+y(\alpha\pm 1)$ are all isotropic. Then (1.1) vanishes for $\beta=\alpha$ and $\beta=\alpha\pm 1$. Setting $\xi=f(y,y)\neq 0$, we have

$$-f(x,x)=\alpha^{J}\xi\alpha=(\alpha\pm 1)^{J}\xi(\alpha\pm 1). \tag{1.2}$$

From this it follows that $\xi=\pm(\alpha^{J}\xi+\xi\alpha)$, which is clearly impossible as $\xi\neq 0$. This proves the lemma.

Let r be a projective semilinear involution in a group of semisimilitudes $U_n(K,f)$ relative to an involutive automorphism σ of K such that $r^2=\epsilon$ and ϵ is in Z. Let K_0 be the subsfield of K left fixed by σ. We must assume that $K_0\neq GF(3)$ in the following lemma. In the applications of this lemma, this is assured, of course, if $\sigma=1$ or if K is not commutative, whence K and K_0 are infinite. In the remaining cases, the sfield K_0 will be identified with the sfield of the projective unitary group which we investigate; so the assumption that this latter sfield is not the field $GF(3)$ is all that will be required in order to apply this lemma.

LEMMA 1.5. *Let S be a nonisotropis subspace of dimension $s\geqq 2$. Assume that the characteristic of K is not 2 and that $K_0\neq GF(3)$. If, for every nonisotropic vector x in S, $r(x)=x\delta_x$ where δ_x is in K, then $r(x)=x\delta$ for every x in S where δ is independent of x. Furthermore $\sigma=\delta^*$.*

It is an immediate consequence of this lemma that if $\sigma=1$, δ is in Z. In all cases σ leaves Z fixed.

Proof. Let x_0 be an arbitrary, nonisotropic vector in S. Then $r(x_0)=x_0\delta_{x_0}$. Let $\delta=\delta_{x_0}$. Let K_1 be the subsfield of K left fixed by $\pi=\delta^*\sigma$. Then $r^2=\epsilon$ implies that $\delta\delta^{\sigma}=\epsilon$, which is in Z. From this it follows that $\pi^2=1$. Then $[K:K_1]=1$ or 2 according as $\pi=1$ or not. This means that $K_1\neq GF(3)$. For, if K is infinite, this is clearly the case; if K is finite, it is commutative and $\pi=\sigma$ whence $K_1=K_0\neq GF(3)$.

Let x_0' be a nonisotropic vector of S orthogonal to x_0. Then $r(x_0')=x_0'\delta_{x_0'}$. Set $\delta=\delta_{x_0'}$. By Lemma 1.4, there exists an element α in K_1 such that $x_0+x_0'\alpha$ is also nonisotropic. Because $K_1\neq GF(3)$, we can choose $\alpha\neq 0$. Then $r(x_0+x_0'\alpha)=(x_6+x_0'\alpha)\delta_{x_0+x_0'\alpha}$. On the other hand $r(x_0+x_0'\alpha)=r(x_0)+r(x_0')\alpha^{\sigma}$. So $x_0(\delta_{x_0+x_0'\alpha}-\delta)=x_0'(\delta'\alpha^{\sigma}-\alpha\delta_{x_0+x_0'\alpha})$. Because x_0 and x_0' are linearly independent relative to K, it follows that $\delta_{x_0+x_0'\alpha}=\delta$ and then that $\delta'\alpha^{\sigma}=\alpha\delta$. But $\alpha^{\pi}=\alpha$ as α is in K_1. So $\alpha^{\sigma}=\alpha^{\delta^*}=\delta^{-1}\alpha\delta$ and $\alpha\delta=\delta\alpha^{\sigma}$. Thus $\delta'=\delta$.

As x_0' was chosen to be an arbitrary nonisotropic vector orthogonal to x_0, $\delta_{x_0'\phi}=\delta$ for ϕ in K. So $r(x_0'\phi)=x_0\delta\phi^{\sigma}=x_0\phi\delta$. Thus $\phi^{\sigma}=\delta^{-1}\phi\delta$ and

$\sigma = \delta^*$. Let $x_0, x_1, \cdots, x_{s-1}$ form an orthogonal nonisotropic basis for S. By the above argument $r(x_i) = x_i\delta$. Then for

$$x = \sum_{i=0}^{s-1} x_i\alpha_i \text{ in } S,\ r(x) = \sum_{i=0}^{s-1} r(x_i)\alpha_i^\sigma = x\delta,$$

which completes the proof of the lemma.

LEMMA 1.6. *Let S be a subspace of E of dimension greater than 1 and let t be an element of $\Gamma U_n(K, f_1)$ relative to an automorphism τ of K. Then if $t(x) = x\delta_x$ with δ_x in K for every vector x of S, $t(x) = x\delta$, where δ is in K and is independent of x in S and τ is the inner automorphism δ^* of K.*

The proof is an obvious simplification of the proof of Lemma 1.5. This lemma includes the case that f_1 is alternating.

II. Involutions of the First Kind.

Henceforth we shall assume that $K \neq GF(3)$.

In this part u^* will represent an involution of the first kind in $PU_n(K, f)$. The usual restrictions on K shall hold. In accordance with Section 1.6, we choose the representative u of u^* which is a p-involution with $p \leqq [n/2]$. Then $E = U^+ \oplus U^-$, where U^+ and U^- are the eigenspaces of u.

LEMMA 2.1. *Assume $n > 2$. If u^* is an extremal involution and v^* an involution commuting with u^*, then a projective involution v in v^* must commute with any projectixe involution u in u^* and v^* is either of the first or third kind.*

That v in v^* must commute with u in u^* is the result at the end of Section 1.6. Thus $v(U^+) = U^+$. If x is in U^+, $v(x) = x\delta_x$. If $v^2 = \beta$, then $\beta = \delta_x{}^2$. If δ_x is in Z, then $\delta_x{}^J\delta_x f(x, x) = f(v(x), v(x)) = f(x, x) \neq 0$. So $\delta_x{}^J\delta_x = 1$. This shows that v and v^* are either of the first or third kind.

2.1. Extremal involutions commuting with involutions of the first kind. Assume $n > 3$. Let u^* be an extremal involution and let $v^* \neq u^*$ be an involution of the first kind commuting with u^*. Let v be the p-involution in v^* with $p \leqq [n/2]$ and with eigenspaces V^+ and V^-. We then have $E = P_1 \oplus P_2 \oplus P_3 \oplus P_4$, where the subspaces P_i are orthogonal and nonisotropic or 0-dimensional. As $\dim U^+ = 1$ one of $P_1 = U^+ \cap V^+$ or $P_2 = U^+ \cap V^-$ contains only the 0 vector. As $u^* \neq v^*$, $u \neq v$, and $U^- \neq V^-$. So $U^- = P_3 \oplus P_4$, where $\dim P_3 > 0$ and $\dim P_4 > 0$. Thus E is decomposed into the direct sum of three orthogonal nonisotropic subspaces U^+, P_3 and P_4.

As $\dim E > 3$, at least one of the P_i is of dimension greater than 1. Let x be a nonisotropic vector in P_j where $\dim P_j > 1$. Let w be the 1-involution with subspaces $W^+ = xK$ and $W^- = (xK)^\perp$; w is in $U_n(K,f)$ by Section 1.6. As W^+ is in P_j, it is left invariant by u and v. Then $W^- = (W^+)^\perp$ is left invariant by u and v, and, consequently, u and v commute with w. That is, w is in $c(u,v)$; so any r in $c(c(u,v))$ must commute with w, as w is a 1-involution. Then $r(x) = x\delta_x$ and this must hold for every nonisotropic vector x in P_j. Consequently by Lemma 1.5, $r(x) = x\delta_j$ with δ_j in Z.

Furthermore, r must commute with u and commute or anticommute with v, as u and v are in $c(u,v)$. However, r cannot anticommute with v, for then would $r(P_1) = P_2$, whereas $\dim P_1 \neq \dim P_2$. So r must leave all P_i invariant. Consequently if $\dim P_k = 1$ and x is in P_k, $r(x) = x\delta_x$. But $r^2(x) = x\epsilon$ for all x in E with ϵ in C. Therefore $\delta_x{}^2 = \delta_j{}^2$. As ϵ can have but two square roots in Z and hence in K, $\delta_x = \pm\delta_j$ and δ_x is in Z. Set $\delta_k = \delta_x$. Then as δ_k is in Z, $r(x) = x\delta_k$ for all x in P_k where $\delta_k = \pm\delta_j$. If $\dim P_k > 1$, the above argument for P_j shows $r(x) = x\delta_k$ for x in P_k and again $\delta_k = \pm\delta_j$. There are thus two choices in sign for $r(x)$ when x is in the two nonzero subspaces $P_k \neq P_j$. There are thus but four possibilities for r, each of which is a representative of 1^* u^*, v^*, or u^*v^*. Therefore $\nu(u^*, v^*) = 4$.

2.2. Extremal involutions commuting with involutions of the third kind. Assume $n > 2$. Let u^* be an extremal involution. Let v^* be an involution of the third kind and let v be in v^*. Then $v^2 = \beta$ is in C_n, where $\beta = \mu^2$ and μ is not in Z. With regard to v, form the group $U_n(K_1, f_1)$ acting on the space E_1 where K_1 is the centralizer of μ in K (cf. Section 1.8). The center of Z_1 of K_1 is $Z(\mu)$. As v_1 is linear, the form f_1 is not alternating (cf. Section 1.8). Then u is represented in $U_n(K_1, f_1)$ by its restriction u_1 to E_1; and in $U_n(K_1, f_1)$, u_1 is a 1-involution with eigenspaces $U_1{}^+ = U^+ \cap E_1$ and $U_1{}^- = U^- \cap E_1$. For the remainder of this section, we assume merely that u_1 is of the first kind in $U_n(K_1, f_1)$. This generalization is needed in Part III.

The K_1-subspaces $U_1{}^+$ and $U_1{}^-$ are orthogonal and nonisotropic (with respect to f_1), at least one of which, say $U_1{}^+$, is of K_1-dimension greater than 1 as $n > 2$. Let x be a nonisotropic vector in $U_1{}^+$ and w_1 be the 1-involution of $U_n(K_1, f_1)$ with subspaces $W_1{}^+ = xK_1$ and $W_1{}^- = (xK_1)^\perp$. Then w commutes with u_1 and its extension w to E is a 1-involution of $U_n(K,f)$ commuting with u and v. Consequently w is in $c(u,v)$ and r in $c(c(u,v))$ must commute with it.

Let θ be the element of K chosen in Section 1.8. so that $\theta^{\mu^*} = -\theta$. Then r or $r' = h_v r$ commutes with $v' = \mu^{-1}v$. Since $h_\theta w = w h_\theta$, r' as well as r commutes with w. Then the restriction r_1 of r or r' to E_1 commutes with w_1. Therefore $r_1(W_1^+) = W_1^+$ and $r_1(x) = x\delta_x$. If r commutes with v', then clearly r_1 is a projective involution of $U_n(K_1, f_1)$ such that $r_1^2 = \epsilon$ where $\epsilon = r^2$. If r anticommutes with v', then r_1 is a projective semilinear involution relative to the automorphism θ^* of K_1 and $r_1^2 = \epsilon\theta^2$. But θ was chosen in Section 1.8 so that $\theta^2 = \alpha$ is in Z. So $r_1^2 = \epsilon\alpha$ is in Z and $(\theta^*)^2 = 1$. In both cases, r_1 satisfies the hypothesis of Lemma 1.5. Consequently $r_1(x) = x\delta$ with δ in K_1 for all x in U_1^+. Furthermore, since θ^* is not an inner automorphism of K_1, r_1 cannot be semilinear with respect to θ^*. Consequently r_1 is a projective involution in $U_n(K_1, f_1)$ and δ is in $Z(\mu)$. As $r_1^2 = \epsilon$, $\delta^2 = \epsilon$, where ϵ is in Z. But then by Section 1.10, we have that δ is in Z or $\delta = \delta_0\mu$ with δ_0 in Z.

If $\dim U_1^- > 1$, the same argument shows that $r(x) = x\delta^-$ with δ^- in $Z(\mu)$ for x in U_1^-. Again $(\delta^-)^2 = \delta^2$; so $\delta^- = \pm\delta$. If $\dim U_1^- = 1$, then as in Section 2.1, $r_1(U_1^-) = U_1^-$ and $r(x) = x\delta_x$ for x in U_1^-. But $\delta_x^2 = \delta^2$; so $\delta_x = \pm\delta$ and is in $Z(\mu)$. This implies that δ_x is independent of x in U_1^-; so we set $\delta_x = \delta^-$.

As r_1 is in $U_n(K_1, f_1)$ and U_1^+ and U_1^- are nonisotropic, $\delta^T\delta = 1$. If δ is in Z, then $\delta^T = \delta^J$. The extension of r_1 to E is r and is seen to be in 1^* or u^* according as $\delta^- = \delta$ or $\delta^- = -\delta$. If $\delta = \delta_0\mu$, by looking at $v_1^{-1}r_1$, we similarly conclude that $\delta_0{}^T = \delta_0{}^J$ and that the extension of $v_1^{-1}r_1$ to E is $v^{-1}r$ and is in 1^* or u^*. So r is in v^* or u^*v^*. This yields

THEOREM 2.1. *If the dimension of E is greater than 3, and u^* is an extremal involution in $PU_n(K, f)$, $\nu_{u^*} = 4$.*

2.3. Nonextremal involutions of the first kind.

THEOREM 2.2. *If the dimension n of E is greater than 4 and not 6, then $\nu_{u^*} > 4$ for nonextremal involutions of the first kind in $PU_n(K, f)$. If $n = 6$, $\nu_{u^*} > 4$ for all 2-involutions u^*; and if, furthermore, E is not commensurable, $\nu_{u^*} > 4$ for all 3-involutions.*

First take $n > 4$ and $n \neq 6$. Let u^* be a p-involution of $PU_n(K, f)$ with $1 < p \leqq [n/2]$ and let u be the p-involution of $U_n(K, f)$ contained in u^*. We shall exhibit an involution $v^* \neq u^*$ commuting with u^* such that $\nu(u^*, v^*) > 4$. Let P_1 be a 1-dimensional nonisotropic subspace of the eigenspace U^+ of u and P_3 a 2-dimensional nonisotropic subspace of the eigenspace

U^- of u. Set $V^+ = P_1 \oplus P_3$, $V^- = (V^+)^\perp$, $P_2 = U^+ \cap V^-$, and $P_4 = U^- \cap V^-$. Then $\dim P_2 = p-1 > 0$ and $\dim P_4 = n-p-2$. Since $p \leqq [n/2]$ and $n \geqq 5$, $\dim P_4 > 0$. So P_1, P_2, P_3, and P_4 are orthogonal nonisotropic subspaces whose direct sum is E. Let v be the 3-involution with positive and negative eigenspaces V^+ and V^-, respectively; then clearly v commutes with u.

Let r_i, $i = 1, 2, 3, 4$, be the involution of $U_n(K, f)$ with eigenspace $R_i^+ = P_i$. Each of u, v, and uv has a positive eigenspace which contains two subspaces P_i; so r_i^* is not one of 1^*, u^*, v^*, or u^*v^*. Now if t is in $c(u, v)$, t must commute with v as $n \neq 6$. If t should anticommute with u, then would $t(P_1) = P_3$, which is impossible as $\dim P_1 \neq \dim P_3$. So t commutes with u and v; hence t leaves invariant P_i, $i = 1, 2, 3, 4$. Consequently t commutes with u and v; thus r_i is in $c(c(u, v))$, which means that $\nu_{u^*} > 4$.

Now let $n = 6$ and u^* be a 2-involution. As above let u be the 2-involution in u^*, and choose P_1, P_2, P_3, and P_4 to be orthogonal nonisotropic subspaces of the eigenspaces U^+ and U^- of u such that $U^+ = P_1 \oplus P_2$, $U^- = P_3 \oplus P_4$ and $\dim P_1 = \dim P_2 = \dim P_3 = 1$. Then $\dim P_4 = 3$. Let v be the 2-involution of $U_n(K, f)$ with eigenspaces $V^+ = P_1 \oplus P_3$ and $V^- = P_2 \oplus P_4$. Clearly $uv = vu$. Define the involutions r_i as above. Then t in $c(u, v)$ must commute with u and v as they are 2-involutions. Hence $t(P_i) = P_i$, $i = 1, 2, 3, 4$; that is, $t(R_i^+) = R_i^+$ and $t(R_i^-) = R_i^-$. Therefore r_i is in $c(c(u, v))$ and $\nu_{u^*} > 4$.

Finally take $n = 6$ and u^* to be a 3-involution. Furthermore, assume E to be not commensurable. Then by Lemma 1.3, there exist nonisotropic 1-dimensional subspaces P_1 and P_4 of U^+ and U^-, respectively, which are not commensurable to each other. Let $P_2 = P_1^\perp \cap U^+$ and $P_3 = P_4^\perp \cap U^-$. Then $E = P_1 \oplus P_2 \oplus P_3 \oplus P_4$ and the 3-involution v with subspaces $V^+ = P_1 \oplus P_3$ and $V^- = P_2 \oplus P_4$ commutes with u. Since $\dim P_1 = \dim P_4 = 1$ and $\dim P_2 = \dim P_3 = 2$, it follows that t in $c(u, v)$ either commutes with both u and v or anticommutes with both of them. However, in the latter case, $t(P_1) = P_4$ in contradiction to the noncommensurability of P_1 and P_4. Thus, t leaves P_i, $i = 1, 2, 3, 4$, invariant, and r_i is in $c(c(u, v))$. So in this case, $\nu_{u^*} > 4$. This proves the theorem.

III. Involutions of the Third Kind.[7]

In this part, u^* will represent an involution of the third kind and u a projective involution contained in u^*. Then $u^2 = \gamma$ in C_n and $\gamma = \lambda^2$ where

[7] It is not necessary to determine whether involutions of the second kind are in $\mathfrak{A}^*$. The argument in Part V allows us to pass over this case.

λ is not in Z. Form the group $\Gamma U_n(K_1, f_1)$ acting on the space $E_1 = U^+$ as in Section 1.8. Since u is linear, f_1 will be hermitian in all cases. Let v be a projective involution of $U_n(K, f)$ which commutes or anticommutes with u and let v_1 be the projective semilinear involution of $\Gamma U_n(K_1, f_1)$ which represents v. We first distinguish four cases when v commutes with u and v_1 is in $U_n(K_1, f_1)$ in accordance as to whether v_1 is of the first, second, third, or fourth kinds. Then we treat the case where v_1 is semilinear with respect to the automorphism θ^* described in Section 1.8.

We have already dealt with the case that v_1 is of the first kind in $U_n(K_1, f_1)$ in Section 2.2.

PROPOSITION 3.1. *Let v be a projective involution in $U_n(K, f)$ commuting with u such that its restriction v_1 to E_1 is a projective involution of the first kind in $U_n(K_1, f_1)$. If $n > 3$, then $\nu(u^*, v^*) = 4$.*

3.1. Commuting projective involutions whose restrictions are of the second kind. Assume $n > 4$. Let v_1 be of the second kind in $U_n(K_1, f_1)$. Then $v_1^2 = \beta$ in C_n, $\beta = \mu^2$ with μ in $Z(\lambda)$ and $\mu^T\mu = -1$, where T is the antiautomorphism defined in Section 1.8. Thus $E_1 = V_1^+ \oplus V_1^-$, where V_1^+ and V_1^- are the totally isotropic eigenspaces of v_1, and $\dim_{K_1} V_1^+ = \dim_{K_1} V_1^-$. Let x be in V_1^+ and pick y in V_1^- such that $f_1(x, y) \neq 0$. Then the plane $W_1^+ \subseteq E_1$ generated by x and y is nonisotropic. Set $W_1^- = (W_1^+)^\perp$ and define w_1 to be the 2-involution of $U_n(K_1, f_1)$ with subspaces W_1^+ and W_1^-. As $W_1^+ = xK_1 \oplus yK_1$, v_1 leaves invariant W_1^+ and consequently W_1^-. Thus w_1 commutes with v_1 and its extension w to E is a 2-involution commuting with u and v. The choice of y in V_1^- is not unique. Indeed, let y'' be in $V_1^- \cap W_1^-$ and set $y' = y + y''$. Then $f_1(x, y') = f_1(x, y) \neq 0$, and y and y' are linearly independent. Let w_1' be the 2-involution of $U_n(K_1, f_1)$ with $W_1'^+ = xK_1 \oplus y'K_1$ and $W_1'^- = (W_1'^+)^\perp$ as its eigenspaces. Then as above w_1' commutes with v_1 and its extension w' to E commutes with u and v.

Now let r be in $c(c(u, v))$. As $n > 4$, r commutes with both w and w'. If r anticommutes with u, we form $r' = h_c r$; r' also commutes with w and w'. In any case, the element r_1 of $\Gamma U_n(K_1, f_1)$ which represents r commutes with w_1 and w_1'. Thus $r_1(W_1^+ \cap W_1'^+) = W_1^+ \cap W_1'^+$; that is, $r(x) = x\delta_x$. Furthermore, as r_1 is linear or semilinear with respect to θ^*, we have by Lemma 1.6 that $r_1(x) = x\delta$ for all x in V_1^+, where δ is in K_1. Furthermore, the possibility that r_1 is semilinear relative to θ^* is eliminated as θ^*, not leaving $Z(\lambda)$ fixed, is not an inner automorphism of K_1. So r_1 must be linear and δ is in $Z(\lambda)$. Similarly $r_1(x) = x\delta^-$ for all x in V_1^-, where δ^- is in K_1.

Now $r_1^2 = \epsilon$ in Z implies that $\delta^2 = (\delta^-)^2 = \epsilon$. Hence $\delta^- = \pm\delta$; and by Section 1.10, δ is in Z or $\delta = \delta_0\lambda$, where δ_0 is in Z. Now

$$f_1(x,y) = f_1(r_1(x), r_1(y)) = \delta^T\delta^- f_1(x,y)$$

for x in V_1^+ and y in V_1^-. As we can further stipulate that $f_1(x,y) \neq 0$, we have that $\delta^T\delta = \pm 1$ according as $\delta^- = \pm\delta$.

If δ is in Z and $\delta^- = \delta$, then $\delta^T = \delta'$ and δ is in C. If δ is in Z and $\delta^- = -\delta$, by replacing r_1 by $v_1 r_1$ in the above argument we obtain that $v_1 r_1(x) = x\delta\mu$ for x in V_1^+ and V_1^- and that $\delta\mu$ is in C. If $\delta = \delta_0\lambda$, we replace r_1 by $u_1 r_1$ or $u_1 v_1 r_1$ to obtain, respectively, that $u_1 r_1(x) = x\delta\lambda$ or $u_1 v_1 r_1(x) = x\delta\mu\lambda$ for x in V_1^+ and V_1^-, where $\delta\lambda$ or $\delta\mu\lambda$ is in C. In any case, we see that there exists an η in C such that ηr_1 coincides with the restriction of 1, u, v, or uv to E_1. Hence r^* is one of 1^*, u^*, v^*, or u^*v^*. Thus we have shown:

PROPOSITION 3.2. *Let v be a projective involution in $U_n(K,f)$ commuting with u whose restriction v_1 to E_1 is a projective involution of the second kind in $U_n(K_1,f_1)$. If $n > 4$, $\nu(u^*, v^*) = 4$.*

3.2. Commuting projective involutions whose restrictions are of the third kind. Assume $n > 2$. Let v be a projective involution of $U_n(K,f)$ commuting with u whose restriction v_1 to E_1 is of the third kind in $U_n(K_1,f_1)$. Then $v_1^2 = \beta$ is in C_n and $\beta = \mu^2$, where μ is not in $Z(\lambda)$ but is in K_1. Let K_2 be the fixed subsfield of K_1 by μ^*; then by Section 1.8, $K_1 = K_2 \oplus K_2\psi$, where $\psi^{\mu^*} = -\psi$ and ψ^2 is in $Z(\lambda)$. As in Section 1.8, consider E_1 to be a vector space E_1^* over K_2. Then $E_1^* = V_1^+ \oplus V_1^-$, where V_1^+ and V_1^- are the eigenspaces of v_1. We form the group $\Gamma U_n(K_2,f_2)$ acting on $E_2 = V_1^+$. Since v_1 is linear in $U_n(K_1,f_1)$, f_2 is always hermitian.

Let K_1' be the fixed subsfield of K by $(\mu^{-1})^* = \mu^*$. Then $K = K_1' \oplus K_1'\psi$. By Section 1.8, the center of K_1' is $Z(\mu)$. As $\mu^{\lambda^*} = \mu$, $\lambda^{\mu^*} = \lambda$; and λ is in K_1'. But $Z(\mu) \neq Z(\lambda)$ as they are both quadratic extensions of Z and μ is not in $Z(\lambda)$. So λ is not in $Z(\mu)$. Hence by Section 1.8, $K_1' = K_2' \oplus K_2'\chi$, where χ is an element of K_1' such that χ^2 is in $Z(\mu)$ and $\chi^{\lambda^*} = -\chi$. Noting that $K_1\theta = K_1\chi$, we shall replace θ by χ. Then we have the following relations:

(i) $K = K_2 \oplus K_2\psi \oplus K_2\chi \oplus K_2\psi\chi$.

(ii) $\lambda^{\chi^*} = -\lambda$; $\mu^{\chi^*} = \mu$; $\lambda^{\psi^*} = \lambda$; $\mu^{\psi^*} = -\mu$.

(iii) χ^2 is in $Z(\mu)$; ψ^2 is in $Z(\lambda)$.

Since v_1 is of the third kind in $U_n(K_1,f_1)$, we have by Section 1.8 that the center of K_2 is $Z(\lambda,\mu)$.

Now let t be in $c(u,v)$. Then if t commutes with u, t is represented in $U_n(K_1,f_1)$ by its restriction to E_1. In accordance with whether t commutes or anticommutes with v, t_1 is represented by a projective semilinear involution t_2 in $\Gamma U_n(K_2,f_2)$ which is linear or semilinear relative to the automorphism ψ^* of K_2. Furthermore, $t^2=\eta$ with η in $Z(\mu)$. If t anticommutes with u, then t is represented in $\Gamma U_n(K_1,f_1)$ by a projective semilinear involution t_1 which is relative to the automorphism χ^* of K_1. In this case, $t_1^2=\eta_1$, where η_1 is in $Z(\mu)$. Then, as $\mu^{\chi^*}=\mu$, t_1 in turn is represented in $\Gamma U_n(K_2,f_2)$ by a projective semilinear involution t_2 which is relative to the automorphism χ^* or $\chi^*\psi^*$ according as t_1 commutes or anticommutes with v_1. Finally $t_2^2=\eta_2$, where η_2 is in $Z(\lambda,\mu)$.

The extension w of a 1-involution w_2 of $U_n(K_2,f_2)$ to E is a 1-involution in $c(u,v)$. Consequently, a projective involution r of $c(c(u,v))$ commutes with w as $n>2$. Therefore, if r_2 represents r in $\Gamma U_n(K_2,f_2)$, r_2 commutes with w_2. This means that r_2 leaves every 1-dimensional nonisotropic space of E_2 invariant. Furthermore, r_2 is a projective semilinear involution relative to an involutive automorphism of K_2 and $r_2^2=\epsilon$ with ϵ in $Z(\lambda,\mu)$. So by Lemma 1.5, $r_2(x)=x\delta$ with δ in K_2 for all x in E_2. Moreover, r_2 must be semilinear with respect to an inner automorphism of K_2. Since χ^*, ψ^* and $\chi^*\psi^*$ are not inner automorphisms, we have that r_2 must be linear. Therefore, r commutes with u and v; so r_2 is the restriction of r to E_2 and ϵ is in Z. Furthermore, δ is in $Z(\lambda,\mu)$ and $\delta^2=\epsilon$. We now apply the result in Section 1.10 twice. First either δ is in $Z(\lambda)$ or $\delta=\delta_1\mu$ with δ_1 in $Z(\lambda)$. Then the square of δ or δ_1, respectively, is an element in Z. This implies that δ is Z, $\delta=\delta_1\mu$, $\delta=\delta_2\lambda$, or $\delta=\delta_3\lambda\mu$ with δ_i in Z. Let u_2 and v_2 be the restrictions of u and v to E_2. Let r_2' be that one of r_2, $u_2^{-1}r_2$, $v_2^{-1}r_2$, or $u_2^{-1}v_2^{-1}r_2$ such that $r_2'(x)=x\delta_i$ for x in E_2 and some $i=1,2,3,4$. Since r_2 is in $U_n(K_2,f_2)$, δ is in Z, and E_2 is nonisotropic, we have $\delta^T\delta=\delta'\delta=1$. Therefore, the extension r' of r_2' to E is in C_n and r is in one of the cosets 1^*, u^*, v^* or u^*v^*. This gives us the following result.

PROPOSITION 3.3. *Let v be a projective involution in $U_n(K,f)$ commuting with u such that its restriction v_2 to E_2 is a projective involution of the third kind in $U_n(K_1,f_1)$. If $n>2$, $\nu(u^*,v^*)=4$.*

3.3. Commuting projective involutions whose restrictions are of the fourth kind. Let v_1 be of the fourth kind in $U_n(K_1,f_1)$. Assume $n>4$. Then $v_1^2=\beta$ with β in C and β not a square in K_1. As in Section 1.9, extend the sfield K_1 to a sfield L_1 by adjoining a square root ρ of β. The

center L_1 is then $Z(\lambda, \rho)$. Then E_1 can be considered as a right vector space F_1 over L_1 by setting $x\rho = v_1(x)$ and f_1 can be extended to a nondegenerate hermitian sesquilinear form g_1 defined on F_1. We then form the group $\Gamma U_m(L_1, g_1)$ acting on F_1.

Let t_1 be the restriction of a projective involution in $U_n(K, f)$ which commutes with u or the representation of a projective involution which anticommutes with u, in which case t_1 is semilinear with respect to the automorphism θ^* of K_1. If, furthermore, t commutes or anticommutes with v, t_1 commutes or anticommutes, respectively, with v_1. Then t_1 can be considered as a projective semilinear involution in $\Gamma U_m(L_1, g_1)$ relative to one of the automorphisms 1, θ^*, σ or $\theta^*\sigma$ where σ is the automorphism of L_1 leaving K_1 fixed and defined by $\sigma : \rho \to -\rho$. It is easily verified that σ and θ^* commute; so all these automorphisms are involutive. The projective involutions u, v, and uv are represented in $U_m(L_1, g_1)$ by the homothetic transformations h_λ, h_ρ, and $h_{\lambda\rho}$, respectively.

Every 1-involution w_1 of $U_m(L_1, g_1)$ extends to a 2-involution w of $U_n(K, f)$ in $c(u, v)$. *If* r_1 in $\Gamma U_m(L_1, g_1)$ represents a projective involution r in $c(c(u, v))$, then r must commute with w as $n > 4$ and r_1 must commute with w_1. So r_1 leaves invariant every 1-dimensional nonisotropic subspace of F_1. Furthermore, r_1 is a projective semilinear transformation relative to an involutive automorphism of L_1 and $r_1^2 = \epsilon$ with ϵ in $Z(\lambda)$. So by Lemma 1.5, $r_1(x) = x\delta$ with δ in L_1 for all x in F_1. However, r_1 cannot be semilinear with respect to the automorphisms, θ^*, σ, or $\theta^*\sigma$ as none of these are inner automorphisms. Consequently, r_1 is linear and must be the representation of a projective involution r commuting with u and v. Therefore, ϵ is in Z and δ is in $Z(\lambda, \rho)$. Also $\delta^2 = \epsilon$. So applying the result of Section 1.10 twice as in Section 3.2, we have that $\delta = \delta_0$, $\delta = \delta_1\rho$, $\delta = \delta_2\lambda$ or $\delta = \delta_3\lambda\rho$, where δ_i is in Z and $\delta_i{}^J\delta_i = 1$. Thus r_1 is the representation of a projective involution in 1^*, u^*, v^*, or u^*v^*. This yields the following result.

PROPOSITION 3.4. *Let v be a projective involution in $U_n(K, f)$ commuting with u such that its restriction v_1 to E_1 is a projective involution of the fourth kind in $U_n(K_1, f_1)$. If $n > 4$, $\nu(u^*, v^*) = 4$.*

Now we consider the case that v anticommutes with u. Then as in Section 1.8, we set $v' = h_\theta v$; v' is a projective semilinear involution in $\Gamma U_n(K, f)$ which commutes with $u' = \lambda^{-1}u$. The restriction v_1 of v' to E_1 is a projective semilinear involution in $\Gamma U_n(K_1, f_1)$ relative to the automorphism θ^* of K_1, and $v_1^2 = \beta$ with β in Z. First note that if $v_1(x) = x\mu$ for any x in E_1, $v_1^2(x) = x\mu\mu^{\theta^*} = x\beta$. Therefore, we distinguish two cases:

Case I, where $\beta = \mu\mu^{\theta^*}$ for some μ in K_1, and Case II, where β is not of the form $\mu\mu^{\theta^*}$ for any μ in K_1.

3.4. Anticommuting projective involutions. Case I. Assume $n > 2$. Let v be a projective involution in $U_n(K,f)$ anticommuting with u such that v is represented by a projective semilinear involution v_1 in $\Gamma U_n(K_1,f_1)$, where ${v_1}^2 = \beta$, and $\beta = \mu\mu^{\theta^*}$ is in Z with μ in K_1. Let K_2 be the fixed subsfield of K_1 by $\theta^*\mu^*$. Since μ is in K_1, $\lambda^{\mu^*} = \lambda$ and $\lambda^{\theta^*\mu^*} = -\lambda$; so $K_1 = K_2 \oplus K_2\lambda$. Let Z_2 be the center of K_2; then an element in Z_2 commutes not only with every element of K_2 but also with λ and therefore with every element of K_1. So $Z_2 \subseteq K_2 \cap Z(\lambda) = Z$. Since Z is left elementwise invariant by $\theta^*\mu^*$, $Z \subseteq Z_2$. Therefore $Z = Z_2$. In the manner of Section 1.8, form the group $\Gamma U_n(K_2,f_2)$ acting on $E_2 = V^+$. The form f_2 will be hermitian or alternating.

Now an element t of $c(u,v)$ which commutes with u and v commutes also with $h_\theta v$ and so can be represented in $U_n(K_2,f_2)$ or $Sp_n(K_2,f_2)$ by the restriction t_2 of t to E_2. If an element t commutes with u and anticommutes with v, then ut will commute with both u and $h_\theta v$ and hence ut can be restricted to a linear transformation in $U_n(K_2,f_2)$ or $Sp_n(K_2,f_2)$. Similarly, if t anticommutes with u and commutes with v, form vt; if t anticommutes with both u and v, form uvt.

In the case f_2 is hermitian, for every 1-dimensional subspace of E_2 which is nonisotropic relative to the form f_2, there exists a 1-involution w_2 in $U_n(K_2,f_2)$ which can be extended to a 1-involution w of $U_n(K,f)$ that commutes with u and v. Denote by r' that one of r, ur, vr, or uvr which commutes with u and v. Then as $n > 2$, r' will commute with the 1-involutions w in $c(u,v)$, and its restriction r_2 to E_2 will commute with the 1-involutions w_2 of $U_n(K_2,f_2)$. Hence by Lemma 1.5, $r_2(x) = x\delta$ with δ in Z for every vector x in E_2. Furthermore $\delta^H\delta = 1$, where H is the antiautomorphism of K_2 corresponding to f_2. As δ is in Z, $\delta^H = \delta^J$. So the extension r' of r_2 to E is a homothetic transformation of E in $U_n(K,f)$. Therefore, r is in one of 1^*, u^*, v^* or u^*v^* and $\nu(u^*,v^*) = 4$.

In the case f_2 is alternating, any plane in E_2 generated by two nonorthogonal vectors will be nonisotropic. So if x is any vector in E_2, pick two linearly independent vectors y and y' not in $(xK)^\perp$ and call W and W' the nonisotropic planes generated by x and y, and x and y', respectively. Then the 2-involutions w_2 and w_2' with W and W' as their respective positive eigenspaces can be extended to 2-involutions w and w' in $c(u,v)$. If r' is the projective involution chosen in the previous paragraph, r' commutes with both w and w'

when $n > 4$ while w and w' are 2-involutions. Then the restriction r_2 of r' to E_2 commutes with w_2 and w_2'; so $r_2(W \cap W') = W \cap W'$. That is, $r_2(x) = x\delta_x$ for each x in E_2. Then by **Lemma 1.5**, $r_2(x) = x\delta$ with δ in Z for all x in E_2. As it was argued in the above paragraph, r is in one of 1^*, u^*, v^*, or u^*v^*.

PROPOSITION 3.5. *Let v be a projective involution of $U_n(K,f)$ anticommuting with u such that v is represented by a projective semilinear involution v_1 of $\Gamma U_n(K_1,f_1)$, where $v_1^2 = \beta$ and $\beta = \mu\mu^{\theta^*}$. If $n > 2$ when f_2 is herimitian and $n > 4$ when f_2 is alternating, $\nu(u^*, v^*) = 4$.*

3.5. Anticommuting projective involutions. Case II. Assume $n > 4$. Now consider the case that v anticommutes with u and is represented by a projective semilinear involution v_1 acting on E_1 such that $v_1^2 = \beta$, where β is not of the form $\mu\mu^{\theta^*}$ for μ in K_1. We then adjoin a square root ρ of β to K_1 forming a sfield $L_1 = K_1(\rho)$ and form the group $\Gamma U_m(L_1, g_1)$ acting on the space F_1 as in Section 1.9. Now ρ satisfies the relation $\phi\rho = \rho\phi^{\theta^*}$ for all ϕ in K_1. In particular if ϕ is in Z_1, the center of L_1, then $\phi^{\theta^*} = \rho^{-1}\phi\rho = \phi$. On the other hand, $\phi(\xi + \rho\eta) = (\xi + \rho\eta)\phi$ for all ξ and η in K_1. This implies that $\phi\xi - \xi\phi = \rho(\eta\phi - \phi\eta)$. So ϕ must commute with ξ and η; hence $Z_1 \subseteq Z(\lambda)$. This implies that $Z_1 = Z$.

Again an element t in $U_n(K,f)$ which commutes with u and v commutes with $h_\theta v$ and hence can be restricted to a projective involution t_1 in $U_n(K_1,f_1)$ which commutes with v_1. Hence t_1 can be considered to be in $U_m(L_1,g_1)$. If t in $c(u,v)$ anticommutes with one or both of u, v, or uv, then one of ut, vt, or uvt will commute with both u and v.

Now let w_1 be an 1-involution of $U_m(L_1,g_1)$. Then w_1 can be extended to a 2-involution w in $U_n(K,f)$ which commutes with u and v. Since $n > 4$, r in $c(c(u,v))$ must commute with w. Denote by r' that one of r, ur, vr, or uvr which commutes with u and v. Then r' commutes with w and is represented by a linear projective involution r_1' of $\Gamma U_m(L_1,g_1)$ commuting with w_1. Hence r_1' leaves invariant every 1-dimensional nonisotropic subspace of F_1. Therefore, by Lemma 1.5 $r_1'(x) = x\delta$ with δ in Z for all x in F_1. Again as in Section 3.4, this implies that $\delta^j\delta = 1$ and thus that r' is in 1^*. Hence r is in one of 1^*, u^*, v^*, or u^*v^*. This yields the following result.

PROPOSITION 3.6. *Let v be a projective involution of $U_n(K,f)$ anticommuting with u such that v is represented by a projective semilinear involution v_1 of $U_n(K_1,f_1)$, where $v_1^2 = \beta$ and β is not of the form $\mu\mu^{\theta^*}$ for μ in K_1. If $n > 4$, $\nu(u^*, v^*) = 4$.*

Now Propositions 3.1 through 3.6 yield the following theorem.

THEOREM 3.1. *If the dimension of E is greater than 4, and if u^* is an involution of the third kind in $PU_n(K,f)$, then $\nu_{u^*} = 4$.*

IV. Involutions of the Fourth Kind.

In this part, u^* will represent an involution of the fourth kind in $PU_n(K,f)$. We shall assume $n > 4$. Pick a representative u in u^*; then $u^2 = \gamma$ in C_n, where γ is not a square in K. As in Section 1.9, form the sfield $L = K(\rho)$, where $\rho^2 = \gamma$. The center of L is $Z(\rho)$. We consider E to be a right vector space F over L by setting $x\rho = u(x)$, and we form the group $\Gamma U_m(L,g)$. Let σ be the automorphism of L leaving K fixed such that $\rho^\sigma = -\rho$. Then an element of $U_n(K,f)$ which commutes or anti-commutes with u can be considered to be an element of $\Gamma U_n(K,f)$ which is, respectively, linear or semilinear relative to the automorphism σ.

We shall designate by v a projective involution commuting or anti-commuting with u. We distinguish various cases and in each show that $\nu(u^*, v^*) = 4$.

4.1. Projective involutions of the first kind in $U_m(L,g)$. Let v be a projective involution of the first kind in $U_m(L,g)$. Then $v^2 = \beta$ with β in C, $\beta = \mu^2$, where μ is in $Z(\rho)$ and $\mu^J\mu = 1$. As in Section 1.6, $F = V^+ \oplus V^-$, where V^+ and V^- are the orthogonal nonisotropic eigenspaces. Since $m > 2$, one of V^+ and V^-, say V^+, must be of L-dimension greater than 1. Let x be a nonisotropic vector in V^+ and w the 1-involution of $U_m(L,g)$ with xK as its 1-dimensional subspace W^+. Since $W^+ \subseteq V^+$ and v is in $U_m(L,g)$, v leaves invariant W^+ and $W^- = (W^+)^\perp$. Consequently, w commutes with v, and thus, w is in $c(u,v)$. Let r be in $c(c(u,v))$. Then as an element of $\Gamma U_m(L,g)$, r is linear or semilinear with respect to the automorphism σ. Furthermore, r must commute with w as $m > 2$; hence $r(x) = x\delta_x$ with δ_x in L for each nonisotropic vector x in V^+. So by Lemma 1.5, $r(x) = x\delta$ with δ in L for x in V^+. Also, r must be semilinear with respect to an inner automorphism of L. But σ does not leave fixed the center $Z(\rho)$ of L and, therefore, is not an inner automorphism of L. So r is linear and δ is in $Z(\rho)$. Also $r^2 = \epsilon$ with ϵ in Z; so $\delta^2 = \epsilon$. It then follows from Section 1.10 that δ is in Z or $\delta = \delta_0\rho$ with δ_0 in Z.

If $\dim_L V^- > 1$, it follows in a similar manner that $r(x) = x\delta^-$ with δ^- in L for x in V^-; and as $r^2 = \epsilon$, we have $(\delta^-)^2 = \delta^2 = \epsilon$. So $\delta^- = \pm\delta$. If $\dim_L V^- = 1$, then $r(x) = x^{\delta_x}$ for x in V^-. Again, ${\delta_x}^2 = \delta^2$; so $\delta_x = \pm\delta$.

So δ_x is in $Z(\rho)$ and hence is independent of x in V^-. Set $\delta_x = \delta^-$. Since g is sesquilinear with respect to the antiautomorphism J extended to L, and V^+ is nonisotropic, we have that $\delta^J\delta = 1$. Thus if δ is in Z, δ is in C. If $\delta = \delta_0\rho$, $\delta_0{}^J\delta_0 = 1$ as $\rho^J = \rho^{-1}$; so δ_0 is in C. In any case r considered an element of $U_n(K,f)$ is in one of the cosets 1^*, u^*, v^*, or u^*v^*. So we have proved the following result.

PROPOSITION 4.1. *Let v be a projective involution of the first kind in $U_m(L,g)$. If $n > 4$, $\nu(u^*,v^*) = 4$.*

4.2. Projective involutions of the second kind in $U_m(L,g)$. Let v be a projective involution of the second kind in $U_m(L,g)$. Then $v^2 = \beta$ with β in C, $\beta = \mu^2$ with μ in $Z(\rho)$, and $\mu^J\mu = -1$. As in Section 1.7, $F = V^+ \oplus V^-$, where V^+ and V^- are totally isotropic subspaces of L-dimension $m/2$.

Let x be an arbitrary vector in V^+ and pick y in V^- such that $g(x,y) \neq 0$. Then $W^+ = xL \oplus yL$ is nonisotropic and the 2-involution w in $U_m(L,g)$ with W^+ and $W^- = (W^+)^\perp$ as its eigenspaces commutes with v. Also $W^- = W^- \cap V^+ \oplus W^- \cap V^-$, where $W^- \cap V^+$ and $W^- \cap V^-$ are totally isotropic subspaces of L-dimension $m/2 - 1$. Since $m > 2$, these subspaces are not zero; so there exists a vector b in $W^- \cap V^-$. Set $y' = y + b$; then $g(x,y') = g(x,y) \neq 0$. Let w' be the 2-involution of $U_m(L,g)$ with $W'^+ = xL \oplus y'L$ and $W'^- = (W')^\perp$ as its eigenspaces. Both w and w' can be considered to be in $c(u,v)$ when considered as transformations of $U_n(K,f)$.

If $n \neq 8$, we can conclude that r in $c(c(u,v))$ commutes with w and w'. We shall show that this is still true when $n = 8$. Then

$$\dim_L W^\pm \cap V^\pm = \dim_L W'^\pm \cap V^\pm = 1.$$

We first assume that r commutes with w; then we show that r commutes with w'. Indeed, if r anticommutes with w', then $r(x)$ is not in $W^+ \cap V^+ = W^+ \cap W'^+$ and so is in $W^+ \cap V^- = yL$ as r must also commute or anticommute with v. Thus $r(x) = y\phi$ with ϕ in L. But also $r(x)$ is in W'^- and so is orthogonal to x. This is a contradiction as $g(x,y) \neq 0$; hence r commutes with w' in this case.

A similar argument shows that if r commutes with w', r commutes with w. So there remains to consider only the case that r anticommutes with w and w'. We shall exclude this case by showing that there exists a third involution w'' in $c(u,v)$ with which r commutes. Then the argument of the previous paragraph will yield an obvious contradiction.

Inasmuch as b is orthogonal to x, y, and y', we have that b is in $W^- \cap W'^- \cap V^-$. It then follows that $r(x) = b\psi$, where ψ is in L. We replace $b\psi$ by b so that $r(x) = b$. Because x and $r(x)$ are in V^+ and V^-, respectively, r anticommutes with v. Thus it follows that there exists a vector a in $V^+ \cap W^-$ such that $r(a) = y$. Set $z = x + a$; then $r(z) = y'$. Also if $\epsilon = r^2$, $r(y) = a\epsilon$ and $r(b) = x\epsilon$.

Of course, it may result that z and $r(z)$ are orthogonal. However, we shall now show that there exists a vector $z_\phi = x + a\phi$ with ϕ in L such that $g(z_\phi, r(z_\phi)) \neq 0$. If τ is the automorphism related to r, then $r(z_\phi) = y\phi^\tau + b$. Setting $\xi = g(x, y) \neq 0$, we have

$$g(z_\phi, r(z_\phi)) = \xi\phi^\tau + \phi^J\xi^J\epsilon. \tag{4.1}$$

If $g(z_\phi, r(z_\phi)) = 0$ for all ϕ in L, we have

$$\phi^J = -\xi\phi^\tau\xi^{-J}\epsilon^{-1}. \tag{4.2}$$

Setting $\phi = 1$ in (4.2), we obtain $\xi^{-J}\epsilon^{-1} = -\xi^{-1}$ and

$$\phi^J = \xi\phi^\tau\xi^{-1} = \phi^{\tau\xi^*}. \tag{4.3}$$

Thus L must be commutative as J is an antiautomorphism. Then it follows from (4.3) that $J = \tau$. There are two possibilities for τ: either $\tau = 1$ or $\tau = \sigma$. In either case, J is the identity on K. Consequently $U_8(K, f)$ is the orthogonal group $O_8(K, f)$. Then by Section 1.9, $\rho = i = \sqrt{-1}$ and $i^J = -i$. Thus $J = \sigma$. But β is in C; so $\beta^J\beta = \beta^2 = 1$. It is easily seen that if $\beta = 1$, then $\mu^J\mu = 1$. If $\beta = -1$, then $\mu = \pm i$; and again $\mu^J\mu = 1$. However, the condition that v is a projective involution of the second kind in $U_m(L, g)$ requires that $\mu^J\mu = -1$. Thus we can conclude that there exists an element ϕ in L such that $g(z_\phi, r(z_\phi)) \neq 0$.

Clearly then the subspace $W''^+ = z_\phi L \oplus r(z_\phi)L$ is nonisotropic. So the 4-involution w'' of $U_8(K, f)$ with W''^+ as its positive eigenspace is in $c(u, v)$. Clearly $r(W''^+) = W''^+$; so r must commute with w''. Therefore, w'' is the involution which we seek.

Consequently when $n > 4$, $r(W^+ \cap W'^+) = W^+ \cap W'^+$; that is, $r(x) = x\delta_x$ for all x in V^+. Then by Lemma 1.6, $r(x) = x\delta$ with δ in L for all x in V^+. Furthermore, r cannot be semilinear relative to σ as σ is not an inner automorphism of L. Therefore, r is linear and δ is in $Z(\rho)$. Furthermore, $\delta^2 = \epsilon$ is in Z, where $\epsilon = r^2$. So by Section 1.10, δ is in Z or $\delta = \delta_0\rho$, where δ_0 is in Z. A similar argument shows that $r(x) = x\delta^-$ with δ^- in $Z(\rho)$ for all x in V^-. We then have $\delta^- = \pm\delta$.

If δ is in Z and $\delta^- = \delta$, r is the homothetic mapping h_δ; and as F is

nonisotropic, $\delta'\delta = 1$. Thus $r^* = 1^*$ in this case. If δ is in Z and $\delta^- = -\delta$, then vr is the homothetic mapping $h_{\mu\delta}$. So in this case $r^* = v^*$. If $\delta = \delta_0\rho$ and $\delta^- = \delta$, then similarly $r^* = u^*$. Finally, if $\delta = \delta_0\rho$ and $\delta^- = -\delta$, then $r^* = u^*v^*$. Thus we have proved:

PROPOSITION 4.2. *Let v be a projective involution of the second kind in $U_m(L, g)$. If $n > 4$, $\nu(u^*, v^*) = 4$.*

4.3. Projective involutions of the third kind in $U_m(L, g)$. Let v be a projective involution of the third kind in $U_m(L, g)$. Then $v^2 = \beta$, $\beta = \mu^2$, where μ is in L but not in $Z(\rho)$. As in Section 1.8, let L_1 be the fixed subsfield of L by μ^* and let θ be an element of L such that $\theta^{\mu^*} = -\theta$ and θ^2 is in $Z(\rho)$. Then $L = L_1 \oplus L_1\theta$. The center of L_1 is $Z(\rho, \mu)$. F can be considered as a right vector space F^* over L_1, and $F^* = V^+ \oplus V^-$. Set $F_1 = V^+$, and take the form g_1 as in Section 1.8. Then form the group $\Gamma U_m(L_1, g_1)$ acting on the space F_1. As v is linear, g_1 is hermitian.

If t is a projective involution of $U_m(L, g)$ which commutes with v, then by Section 1.8, t can be restricted to a projective involution t_1 of $U_m(L_1, g_1)$. Conversely, if t_1 is a projective involution of $U_m(L_1, g_1)$, t_1 can be extended to a projective involution t of $U_m(L, g)$. If t anticommutes with v, then t is represented by a projective semilinear involution t_1 of $\Gamma U_m(L_1, g_1)$ which is semilinear relative to the automorphism θ^* of L_1.

The situation is not quite so simple if t is a projective semilinear involution of $\Gamma U_m(L, g)$ relative to the automorphism σ. If we can show that $\mu^\sigma = \mu$, then by Section 1.8 we can restrict t to a semilinear projective involution t_1 of $\Gamma U_m(L_1, g_1)$ which is semilinear with respect to the automorphism σ. Indeed, as we shall see, this is true if t is in $c(c(u, v))$ and t commutes with v. Furthermore, if t is such a projective semilinear involution in $c(c(u, v))$ which anticommutes with v, then $\mu^\sigma = -\mu$; but then we show that t has a restriction t_1 to F_1 which is a projective semilinear involution of $\Gamma U_m(L_1, g_1)$ relative to the automorphism σ of L.

Therefore, let t be a projective semilinear involution of $\Gamma U_m(L, g)$ relative to the automorphism σ of L which is in $c(c(u, v))$. Let x be a vector of F_1 which is nonisotropic with respect to the form g_1. Set $W_1^+ = xL_1$ and $W_1^- = (W_1^+)^\perp$, and let w_1 be the 1-involution of $U_m(L_1, g_1)$ with W_1^+ and W_1^- as its eigenspaces. As $m > 2$, a projective semilinear involution r in $c(c(u, v))$ must commute with the extension w of w_1 to F. In particular, t commutes with w. As a transformation of F, w has positive eigenspace $W^+ = xL_1 \oplus xL_1\theta = xL$. Therefore $t(x) = x\delta_x$ with δ_x in L. Set $\delta_x = \delta$.

Since $t^2 = \epsilon$ in Z, we have $\delta\delta^\sigma = \epsilon$. Then it is easily verified that $\delta^*\sigma = \sigma(\delta^\sigma)^*$; so the automorphism $\pi = \delta^*\sigma$ is an involution. Let L_0 be the fixed subsfield of L by π. Then L_0 is of index 2 in L. Furthermore, $\rho^\pi = \rho^{\delta^*\sigma} = \rho^\sigma = -\rho$. So $L = L_0 \oplus L_0\rho$.

We next show that $L_0 = K$. Let ϕ be in K; then $\phi = \phi_1 + \rho\phi_2$, where ϕ_1 and ϕ_2 are in L_0. But $\phi^\sigma = \phi$; so $\phi_1{}^\sigma - \rho\phi_2{}^\sigma = \phi_1 + \rho\phi_2$. This implies that $\phi_1{}^\sigma = \phi_1$ and $\phi_2{}^\sigma = -\phi_2$. So ϕ_1 is in K and $\phi_2 = \rho\phi_3$ with ϕ_3 in K. Thus $\phi = \phi_1 + \phi_3\gamma = \phi_1 + \rho\phi_2$, and the representation of ϕ is unique only if $\phi_2 = 0$ and $L_0 = K$.

But then the restriction of π to $L_0 = K$ coincides with the restriction of δ^* to L_0. But also π is the identity on L_0; so δ commutes with every element of K. Also δ will commute with every element of $K\rho$; so δ is in $Z(\rho)$. Now $t(v(x)) = t(x)\mu^\sigma = x\delta\mu^\sigma$ and $v(t(x)) = v(x)\delta = x\mu\delta$. So if t commutes with v, $\mu^\sigma = \mu$ and the method of Section 1.8 applies and we can restrict t to F_1.

If t anticommutes with v, then $\mu^\sigma = -\mu$. Therefore, $\mu^*\sigma = \sigma\mu^*$ and σ leaves L_1 invariant. Furthermore, t also anticommutes with $h_{\mu^{-1}}$; so t commutes with $v' = h_{\mu^{-1}}v$ and consequently can be restricted to a semilinear projective involution of $\Gamma U_m(L_1, g_1)$ acting on F_1 relative to the automorphism σ of L_1.

So in general, a projective semilinear involution r in $c(c(u, v))$ can be restricted to or is represented by a projective semilinear involution r_1 in $\Gamma U_m(L_1, g_1)$ which is linear or semilinear with respect to the involutive automorphism θ^* or σ of L_1. Since the choice of a nonisotropic vector x in F_1 was arbitrary, r_1 must commute with every 1-involution in $U_m(L_1, g_1)$. Hence r_1 leaves invariant every 1-dimensional nonisotropic subspace of F_1. Since $r_1{}^2 = \epsilon$ with ϵ in $Z(\rho, \mu)$, Lemma 1.5 yields that $r(x) = x\delta$ with δ in L_1 for all x in F_1. Furthermore, r_1 must be semilinear with respect to an inner automorphism of L_1. However, neither θ^* nor σ is an inner automorphism. Consequently, r_1 must be linear; δ is in $Z(\rho, \mu)$; and $\delta^2 = \epsilon$ with ϵ in Z. Then, as it was argued in Section 3.3, r is in one of 1^*, u^*, v^*, or u^*v^*. This yields the following result.

PROPOSITION 4.3. *Let v be a projective involution of the third kind in $U_m(L, g)$. If $n > 4$, $\nu(u^*, v^*) = 4$.*

4.4. Projective involutions of the fourth kind in $U_m(L, g)$. Let v be a projective involution of the fourth kind in $U_m(L, g)$. Then $v^2 = \beta$ in C_n and β is not a square in L. We adjoin a square root ρ' of β to L forming

the sfield $M = L(\rho') = K(\rho, \rho')$. The center of L is $Z(\rho, \rho')$. As in Section 1.9, the space F can be considered as a right vector space G over M of dimension $q = m/2 = n/4$. The form g can be extended to a nondegenerate hermitian sesquilinear form h defined on G relative to the extension of the antimorphism J to M. We then form the group $\Gamma U_q(M, h)$. An element in $U_n(K, f)$ which commutes with u and v can be considered to be in $U_q(M, h)$ and conversely. An element in $U_n(K, f)$ which anticommutes with u and commutes with v, or which commutes with u and anticommutes with v, or which anticommutes with both u and v can be considered to be in $\Gamma U_q(M, h)$ relative to the automorphisms σ, σ', or $\sigma\sigma'$, respectively, where $\sigma: \rho \to -\rho$, $\sigma: \rho' \to \rho'$, $\sigma': \rho' \to -\rho'$, $\sigma': \rho \to \rho$, and σ and σ' leave K fixed.

Now let w be a 1-involution in $U_q(M, h)$; w can be considered as a 4-involution in $U_n(K, f)$ which commutes with both u and v. If r is in $c(c(u, v))$ then r commutes or anticommutes with w. If $n > 8$, we can conclude immediately that r commutes with w. Let $n = 8$ and assume r anticommutes with w; then r, as an element of $\Gamma U_q(M, h)$, is such that $r(x) = y$, where x is in W^+, the positive eigenspace of w, and where y is in W^-, the negative eigenspace of w. The vector y is orthogonal to x and is nonisotropic as $h(y, y) = h(x, x) \neq 0$. Also $r^2 = \epsilon$ with ϵ in Z, so $r(y) = r^2(x) = x\epsilon$. By Lemma 1.4, for every element ϕ of M, there exists an element $\psi = \phi$, $\phi + 1$, or $\phi - 1$ such that $z_\psi = x + y\psi$ is nonisotropic; and r commutes or anticommutes with the 1-involution w_ψ in $U_n(M, h)$ with $z_\psi M$ as its positive eigenspace. In the first case $r(z_\psi) = z_\psi \delta$. This yields $y + x\psi^\tau\epsilon = x\delta_\psi + y\psi\delta_\psi$ where $\tau = 1, \sigma, \sigma'$, or $\sigma\sigma'$ is the automorphism of M associated with r. Due to the linear independence of x and y, we obtain first that $\delta_\psi = \psi^{-1}$ and then that $\psi^\tau\psi = \epsilon^{-1}$, which is in Z. If r anticommutes with w_ψ, then $r(z_\psi)$ is orthogonal to z_ψ. Hence $h(z_\psi, r(z_\psi)) = \xi\psi^\tau\epsilon + \psi^J\xi = 0$ where $\xi = h(x, x) = h(y, y) \neq 0$. This yields $\psi^J = -\xi\psi^\tau\xi^{-1}\epsilon$.

Now by the proper choice of ϕ, we shall obtain contradictions to two of the equations of the above paragraph and thus conclude that r commutes with w. We first note that $Z \neq GF(3)$; otherwise $\beta = \pm 1$ and $\gamma = \pm 1$. This is impossible inasmuch as neither β nor γ can be equal to 1 without u and v being of the first kind in $U_n(K, f)$ while if $\beta = \gamma = -1$, then v would be of the first kind in $U_m(L, g)$. So we choose $\omega_0 \neq 0, \pm 1$ in Z and take $\phi = \rho + \rho' + \omega_0$. Then $\psi = \rho + \rho' + \omega$ where $\omega \neq 0$ and is in Z. Computation shows that

$$(4.4) \quad \psi^\tau\psi = (\omega^2 + \rho^\tau\rho + \rho'^\tau\rho') + (\omega\rho + \omega\rho^\tau) + (\omega\rho' + \omega\rho'^\tau) + (\rho^\tau\rho' + \rho'^\tau\rho)$$

where the terms have been collected to give the components of $\psi^\tau\psi$ according to the direct decomposition $M = K \oplus K\rho \oplus K\rho' \oplus K\rho\rho'$. Now if r is to commute with w_ψ, then $\psi^\tau\psi$ must be in Z. Hence the last three components in (4.4) vanish. As $\omega \neq 0$, $\rho^\tau = -\rho$ and $\rho'^\tau = -\rho'$. But then $(\rho^\tau\rho' + \rho'^\tau\rho) \neq 0$. Thus $\psi^\tau\psi$ is not in Z and r cannot commute with w_ψ.

If r anticommutes with w_ψ, we have shown that $\psi^J = -\xi\psi^\tau\xi^{-1}\epsilon$. Since ψ is in $Z(\rho, \rho')$, $\xi\psi\xi^{-1} = \psi$. Also $\rho^J = \rho^{-1} = \rho\gamma^{-1}$ and $\rho'^J = \rho'^{-1} = \rho'\beta^{-1}$. So we now have

$$\rho\gamma^{-1} + \rho'\beta^{-1} + \omega^J = -(\rho^\tau\epsilon + \rho'^\tau\epsilon + \omega^\tau\epsilon). \tag{4.5}$$

The antiautomorphism J when restricted to Z is either the identity automorphism or an automorphism of order 2. Let Z_0 be the fixed subfield of Z by J; then either $Z = Z_0$ or $Z = Z_0 + Z_0\chi$ where $\chi^J = -\chi$. In the first case, $\omega^J = \omega$. Furthermore $\omega^\tau = \omega$ as ω is in K; so from (4.5), $\epsilon = -1$. In the second case, $\omega = \omega_1 + \omega_2\chi$ where ω_1 and ω_2 are in Z_0. Now (4.5) yields $\epsilon = \pm 1$. Also from (4.5), we have that $\rho\gamma^{-1} = -\rho^\tau\epsilon$ and $\rho'\beta^{-1} = -\rho'^\tau\epsilon$. So if $\tau = 1$ or $\tau = \sigma\sigma'$, then $\beta = \gamma = \pm\epsilon$; this is impossible as γ is a square in L while β is not. If $\tau = \sigma$ or $\tau = \sigma'$, then $\beta = -\gamma = \pm\epsilon$. But $\epsilon = \pm 1$; so one of β or γ equals 1, which is a square in K, contrary to the choice of u and v. This means that r cannot anticommute with w_ψ, which is in $c(u, v)$. As we have also shown that r cannot commute with w_ψ, the assumption that r anticommutes with w yields a contradiction.

This means that r commutes with every 1-involution in $U_q(M, h)$ and hence leaves every 1-dimensional nonisotropic subspace of G invariant. Then by Lemma 1.5, $r(x) = x\delta$ with δ in M for all x in G. If r is linear, then δ is in $Z(\rho, \rho')$. If r is semilinear with respect to an automorphism $\pi = \sigma$, σ', or $\sigma\sigma'$ of M, then π, to be an inner automorphism, must leave $Z(\rho, \rho')$ elementwise invariant, which it does not. Hence δ is in $Z(\rho, \rho')$ and $\delta^2 = \epsilon$ in Z. Again applying the result of Section 1.10 twice, we obtain that $\delta = \delta_0$, $\delta = \delta_1\rho$, $\delta = \delta_2\rho'$, or $\delta = \delta_3\rho\rho'$ with δ_i in Z. As r is a homothetic mapping in $U_q(M, h)$, it follows that $\delta^J\delta = 1$ and hence that $\delta_i^J\delta_i = 1$. Then one of r, ur, vr, or uvr is a homothetic transformation h_η in $U_q(M, h)$, where η is in C. So h_η can be considered in C_n and r is in one of the cosets 1^*, u^*, v^*, or u^*v^*. Therefore, we have the following result.

PROPOSITION 4.4. *Let v be a projective involution of the fourth kind in $U_m(L, g)$. If $n > 4$, $\nu(u^*, v^*) = 4$.*

Now we treat the case that v considered as a projective involution of

$U_n(K,f)$ anticommutes with u. Then v considered as an element of $\Gamma U_m(L,g)$ is a projective semilinear involution relative to the automorphism σ of L. If $v^2=\beta$ where β is in C and if there exists a vector x in F such that $v(x)=x\mu$, then $v^2(x)=x\mu\mu^\sigma$. We therefore distinguish the two cases: β is of the form $\mu\mu^\sigma$ with μ in L (Case I) and β is not of the form $\mu\mu^\sigma$ with μ in L (Case II).

4.5. Projective semilinear involutions in $\Gamma U_m(L,g)$. Case I. Let v be a projective semilinear involution in $\Gamma U_m(L,g)$ relative to the automorphism σ of L such that $v^2=\beta$ and $\beta=\mu\mu^\sigma$ with μ in L. Then as in Section 1.8, form the subsfield L_1 of L consisting of the elements of L left fixed by the automorphism $\sigma'=(\mu^{-1})^*\sigma$. $L=L_1\oplus L_1\theta$, where θ is an arbitrary element of L such that $\theta^{\sigma'}=-\theta$. In particular, we pick $\theta=\rho$. By the argument in Section 1.8, it follows that the center Z_1 of L_1 is contained in $Z(\rho)\cap L_1$. But also $Z\subseteq Z_1$, and as ρ is not in L_1, it follows that $Z_1=Z$.

Now F when considered as a L_1-vector space decomposes into the direct sum $V^+\oplus V^-$ of the eigenspaces of v. Set $F_1=V^+$ and form the group $\Gamma U_m(L_1,g_1)$ acting on F_1 as in Section 1.8. An element t which commutes with both u and v can be restricted to be in $U_m(L_1,g_1)$ or $Sp_m(L_1,g_1)$, and conversely. If t anticommutes with u and commutes with v, then vt commutes with both u and v and so can be restricted to F_1. Similarly in the other cases ut or uvt can be restricted to F_1.

In case g_1 is herimitian, let w_1 be a 1-involution in $U_m(L_1,g_1)$. Then the extension w of w_1 to E is then a 2-involution in $c(u,v)$. Then if r is in $c(c(u,v))$ and r' is that one of r, ur, vr, or uvr which commutes with u and v, r and r' commute with w. The restriction r_1 of r' to F_1 commutes with w_1. Thus it follows that r_1 leaves invariant every 1-dimensional subspace of F_1.

In case g_1 is alternating, let x be a vector of F_1. Then x lies in a maximal totally isotropic subspace v; and, corressponding to V, there exists a complementary totally isotropic subspace V' (cf. Dieudonné [4; p. 6]). Now it is clear that we can apply the argument in Section 4.2 to conclude that $r_1(x)=x\delta_x$.

Consequently, by Lemma 1.5 or Lemma 1.6, we can conclude that $r_1(x)=x\delta$ with δ in Z for all x in F_1. Furthermore, as F_1 is nonisotropic relative to g_1, $\delta^T\delta=1$, where T is the antiautomorphism of L_1 corresponding to g_1. Since δ is in Z, $\delta^T=\delta^J$; so δ is in C and the extension r' of r_1 to E is in C_n. Thus r is in one of 1^*, u^*, v^*, or u^*v^*.

PROPOSITION 4.5. *Let v be a projective semilinear involution in*

$\Gamma U_m(L, g)$ *relative to the automorphism* σ *of* L *such that* $v^2 = \beta$ *and* β *is of the form* $\mu\mu^\sigma$ *with* μ *in* L. *If* $n > 4$, *then* $\nu(u^*, v^*) = 4$.

4.6. Projective semilinear involutions in $\Gamma U_m(L, g)$. Case II. Let v be a projective semilinear involution of $\Gamma U_m(L, g)$ relative to the automorphism σ of L such that $v^2 = \beta$, where β is not of the form $\beta = \mu\mu^\sigma$ with μ in L. Then, as in Section 1.9, adjoin a square root ρ' of β to L forming the sfield $M = L(\rho') = K(\rho, \rho')$. If ϕ is in L, the relation $\phi\rho' = \rho'\phi^\sigma$ holds; in particular $\rho\rho' = -\rho'\rho$. Let Z_M be the center of M. Then clearly $Z \subseteq Z_M \subseteq Z(\rho)$ as $Z(\rho)$ is the center of L. Furthermore, ρ is not in Z_M; and since $[Z(\rho) : Z] = 2$, $Z_M = Z$.

Now form the M-space G by setting $x\rho' = v(x)$ and extend the anti-automorphism J to M by setting $\rho'^J = \rho'^{-1}$. Then, as in Section 1.9, extend the form g to a hermitian sesquilinear form h defined on G and form the group $\Gamma U_q(M, h)$, where $q = m/2 = n/4$. An element t of $U_n(K, f)$ which commutes with both u and v can be considered to be in $U_q(M, h)$ and conversely. In the other cases, ut, vt, or uvt will commute with both u and v; and so one of them can be considered to be in $U_q(M, h)$.

A 1-involution w_1 in $U_q(M, h)$ can then be extended to a 4-involution w in $U_n(K, f)$ which commutes with u and v. Then r in $c(c(u, v))$ commutes or anticommutes with w. If $n > 8$, it follows immediately that r must commute with w. If $n = 8$, the argument of Section 4.4 shows that r must commute with w.[8] Let r' be that one of r, ur, vr, or uvr which commutes with both u and v. Then r' commutes with w and its restriction r_1 to G is a linear projective involution commuting with w_1. That is, r_1 commutes with every 1-involution of $U_q(M, h)$ and so must leave invariant every 1-dimensional nonisotropic subspace of G. Then by Lemma 1.5, $r_1(x) = x\delta$ with δ in Z for every vector x of G. Since r_1 is in $U_q(M, h)$, $\delta^J\delta = 1$ and the extension r' of r_1 to E is a homothetic mapping of E. Thus r' is in 1^*, and r is in one of 1^*, u^*, v^*, or u^*v^*.

---

[8] The argument of Section 4.4 should be altered at two points. First, it should be remarked that τ can be taken to be the identity automorphism, for we can replace r by that one of r, ur, vr, or uvr which commutes with u and v. The extension of this transformation to $U_q(M, h)$ is linear. Second, we should remark that it is possible to choose $\xi \neq \pm 1$ by replacing, if necessary, ξ by $\alpha^J\xi\alpha = h(x\alpha, x\alpha)$ where $\alpha = \rho + \rho'$ or $\alpha = 1 + \rho + \rho'$. Then take $\phi = \xi$ so that $\psi = \xi$ or $\xi \pm 1$. Hence $\psi \neq 0$ and $\psi^J = \psi$. If $\psi^2 = \epsilon^{-1}$, it follows that $\psi^2 = \pm 1$ as $\epsilon^J\epsilon = 1$. But we can assume that $\psi^2 \neq 1$ by a replacement of ξ by $\alpha^J\xi\alpha$ if necessary. So $\epsilon = \epsilon^{-1} = \psi^2 = -1$. If $\psi^2 \neq \epsilon^{-1}$, then $\psi^J = \psi = -\xi\psi\xi^{-1}\epsilon = -\psi\epsilon$. So again $\epsilon = -1$. Now choose $\phi = \rho$. Clearly $\psi^2 \neq -1$; so $\psi^J = \xi\psi\xi^{-1}$. This means that $\rho^J = \xi\rho'\xi^{-1}$. Similarly $\rho^J = \xi\rho'\xi^{-1}$ and $(\rho\rho')^J = \xi\rho\rho'\xi^{-1}$. This implies that $(\rho\rho')^J = \rho^J\rho'^J = (\rho'\rho)^J$ and thus that $\rho\rho' = \rho'\rho$, which is a contradiction.

PROPOSITION 4.6. *Let v be a projective semilinear involution in $\Gamma U_m(L, g)$ relative to the automorphism σ of L such that $v^2 = \beta$, where β is not of the form $\beta = \mu\mu^\sigma$ with μ in L. If $n > 4$, then $\nu(u^*, v^*) = 4$.*

Propositions 4.1 through 4.6 yield the following Theorem.

THEOREM 4.1. *If the dimension n of E is greater than 4 and if u^* is an involution of the fourth kind in $PU_n(K, f)$, then $\nu_{u^*} = 4$.*

V. Characterization of Extremal Involutions.

We now characterize the extremal involutions of $PU_n(K, f)$ when $n > 4$ principally by the group-theoretical invariant ω_{u^*}. When $n = 6$ and E is commensurable, we use other methods.

5.1. Extremal involutions.

THEOREM 5.1. *If the dimension n of E is greater than 4 and E is not commensurable when $n = 6$, then $\omega_{u^*} = 8$ for extremal involutions u^* of $PU_n(K, f)$.*

Let u be the 1-involution in u^*. Let v^* and w^* be the distinct involutions of $\mathfrak{A}^*$ commuting with u^* such that $v^*w^* = w^*v^* \neq u^*$, $v^* \neq u^*$, and $w^* \neq u^*$. Let v and w be in v^* and w^*, respectively. By Lemma 2.1, it follows that every projective involution in $c(u, v, w)$ and $c(c(u, v, w))$ is of the first or third kind and commutes with u. By Theorems 2.1 and 2.2, we have that v^* or w^* is an extremal involution if it is of the first kind.

Any element t in $U_n(K, f)$ which commutes with u can be restricted to a transformation t_0 of U^-, the $(n-1)$-dimensional eigenspace of u. As U^- is nonisotropic, the form f can be restricted to a nondegenerate hermitian sesquilinear form f_0 on U^-. The group of restrictions to U^- of elements of $U_n(K, f)$ is then the group $U_{n-1}(K, f_0)$ acting on U^-. The sets $c(u, v, w)$ and $c(c(u, v, w))$ reduce to the sets $c(v_0, w_0)$ and $c(c(v_0, w_0))$ of projective involutions of $U_{n-1}(K, f_0)$. As elements of $U_{n-1}(K, f_0)$, v_0 and w_0 are the same kind of involutions that v and w are. When $n - 1 > 4$, it follows from Theorems 2.1 and 3.1 that $c(c(v_0, w_0))$ contains representatives from only the four cosets 1^*, v_0^*, w_0^*, and $v_0^* w_0^*$ of $PU_{n-1}(K, f_0)$.

A closer investigation shows that this is even true when $n - 1 = 4$. Indeed, this follows from Theorem 2.1 if v_0^* or w_0^* is an extremal involution. If v_0^* and v^* are of the third kind, then $v^2 = \beta$ and $v_0^2 = \beta$, where $\beta = \mu^2$ and μ is not in Z. By Section 1.8, E, considered as a right vector space

over K_1, the centralizer of μ in K, is the direct sum of the K-subspaces V^+ and V^-, where $v(x) = \pm x\mu$ for x in V^+ or V^- respectively. Similarly V^-, considered as a K_1-space, is the direct sum $V_0^+ \oplus V_0^-$, where $v_0(x) = \pm x\mu$ for x in V_0^+ or V_0^-, respectively. As in Section 1.8, we form the group $U_n(K_1, f_1)$ acting on the subspace $E_1 = V^+$ and $U_{n-1}(K_1, f_1')$ acting on V_0^+. The form f_1' is the restriction of f_1 to $E_1' = V_0^+$ since

$$E_1 = U^+ \cap V^+ \oplus U^- \cap V^+ = U^+ \cap V^+ \oplus E_1'.$$

If w commutes with v, then by Lemma 2.1, w restricted to E_1 must commute with the restriction u_1 of u to E_1 as u_1 is a 1-involution in $U_n(K_1, f_1)$. Furthermore, then w is of the first or third kinds in $U_n(K_1, f_1)$. So it follows that w_0 is of the first or third kinds in $U_{n-1}(K_1, f_1')$. Hence by Propositions 3.1 and 3.3, $c(c(v_0, w_0))$ again contains projective involutions from only four cosets.

Finally if w anticommutes with v, it is represented by a projective semilinear involution w' in $\Gamma U_n(K_1, f_1)$ relative to an automorphism θ^* of K_1, where θ is an element of K such that $\theta^{\mu^*} = -\theta$. Then as w' commutes with u_1, $w'(U^+ \cap E_1) = U^+ \cap E_1$; so $w(x) = x\lambda$ for some x in $U^+ \cap E_1$. This means that $w'^2 = \gamma$, where $\gamma = \lambda\lambda^{\theta^*}$.

As in Sections 1.8 and 3.4, E_1, when considered as a right vector space over K_2, the fixed subsfield of K_1 by $\theta^*\lambda^*$, is the direct sum of the eigenspaces W^+ and W^- of w'. Since w' commutes with u_1 and $U^+ \cap E_1$ is a K_1-subspace, it follows that $U^+ \cap E_1 = U^+ \cap W^+ \oplus U^+ \cap W^-$. If x is in $U^+ \cap W^+$, certainly $f(x, x) \neq 0$. But then, if f_2 is the sesquilinear form defined on W^+ and derived from f_1 in the manner of Sections 1.8 and 3.4, f_2 is but a scalar multiple of the form f restricted to W^+. So $f_2(x, x) \neq 0$. This means that f_2 is not an alternating form. Hence the form f_2' derived from f_1' as in Sections 1.8 and 3.4 is also not alternating inasmuch as f_2' is the restriction of f_2 to $W^+ \cap U^-$. Then by Proposition 3.5 in this last case the projective involutions of $c(c(v_0, w_0))$ come from only the cosets 1^*, v_0^*, w_0^*, and $v_0^* w_0^*$.

Let r^* be an involution of $c(c(u^*, v^*, w^*))$. Then r in r^* can be chosen so that its restriction to U^- is one of 1_0, v_0, w_0, or $v_0 w_0$. Denote by r' that one of 1, v, w, or vw whose restriction to U^- coincides with that of r. Then $r'^2 = r^2 = \epsilon$ in C_n. Furthermore $\epsilon = \delta^2$ where δ is in Z', the center of K', a subsfield of K of index 1 or 2 according as r is of the first or third kind. Let R^+ and R^- be the K'-subspaces of r and R'^+ and R'^- be the K'-subspaces of r'. Then since r and r' coincide on U^-, we have $U^- \cap R^+ = U^- \cap R'^+$ and $U^- \cap R^- = U^- \cap R'^-$. Furthermore

$$U^+ = U^+ \cap R^+ \oplus U^+ \cap R^- = U^+ \cap R'^+ \oplus U^+ \cap R'^-.$$

One of these summands may be zero. Indeed, if r and r' are of the first kind, then $U^+ = U^+ \cap R^+$ or $U^+ = U^+ \cap R^-$ and also $U^+ = U^+ \cap R'^+$ or $U^+ = U^+ \cap R'^-$. Therefore for x in U^+, $r(x) = \pm x\delta$ and $r'(x) = \pm x\delta$. Thus there are but two possibilities for $r(x)$ when x is in $U^+ : r(x) = r'(x)$ or $r(x) = -r'(x)$.

If r and r' are of the third kind, then $\dim_{K'} U^+ \cap R'^{\pm} = 1$. Since r is in $c(c(u, v, w))$ while r' is in $c(u, v, w)$, r and r' must commute or anticommute. As they coincide in U^-, they commute. Therefore, $U^+ \cap R^+$ is equal to $U^+ \cap R'^+$, or $U^+ \cap R'^-$, and $U^+ \cap R^-$ is equal to $U^+ \cap R'^-$ or $U^+ \cap R'^+$, respectively. That is, for x in $U^+ = U^+ \cap R^+ \oplus U^+ \cap R^-$, $r(x) = r'(x)$ or $r(x) = -r'(x)$. Consequently 1_0^*, v_0^*, w_0^*, and $v_0^* w_0^*$ are the restrictions of at most 8 involutions in $c(c(u^*, v^*, w^*))$. Since $\omega(u^*, v^*, w^*) \geqq 8$, we have proven the Theorem 5.1.

We shall denote by $\mathcal{J}^*$ the set of involutions of $\mathfrak{I}^*$ for which $\omega_{u^*} = 8$.

5.2. Involutions of the second kind. We assume that $n > 4$. It can be shown that the involutions of the second kind are not in $\mathcal{J}^*$ providing $n > 6$. To treat the case $n = 6$, we consider the set $\mathcal{K}^*$ of involutions of $PU_n(K, f)$ each of which is not conjugate to the product of itself with two other distinct commuting involutions with which it commutes and to which it is conjugate.

THEOREM 5.2. *Let the dimension of E be greater than 4. Then the set $\mathcal{K}^*$ contains the extremal involutions of $PU_n(K, f)$ but not the involutions of the second kind.*

Let u^*, v^*, w^* be distinct commuting and conjugate extremal involutions. Then by Section 1.6, $u^* v^* w^*$ is a 3-involution and hence not conjugate to an extremal involution.

Let u^* be an involution of the second kind in $PU_n(K, f)$. Pick u to be an involution of $U_n(K, f)$ in u^* with totally isotropic eigenspaces U^+ and U^- of dimension $m = n/2$. Dieudonné [2; p. 7] has shown that if U_0^+ and U_0^- are any two maximal totally isotropic subspaces whose sum $U_0^+ \oplus U_0^- = E$, then there exists an element t of $U_n(K, f)$ mapping U^+ onto U_0^+ and U^- onto U_0^-. The proof of Dieudonné is stated for the symplectic groups but applies in this case as well. Thus if $u^2 = \gamma$, where $\gamma = \lambda^2$, we define u_0 such that $u_0(x) = \pm x\lambda$ according as x is in U_0^+ or U_0^-. Then u and u_0 are conjugate in $U_n(K, f)$. Since every involution u_0^* of the second kind in $PU_n(K, f)$ contains an involution u_0 as defined above, u^* is conjugate to the involution u_0^*.

Now pick a basis $x_1, x_2, \cdots, x_m$ for U^+ and $x_{m+1}, x_{m+2}, \cdots, x_n$ for U^- such that $f(x_i, x_j) = 0$ except when $j - i = \pm m$, in which case $f(x_i, x_j) = 1$. Let $P_1 = x_1K$, $P_2 = x_2K$, $P_3 = x_3K \oplus \cdots \oplus x_mK$, $P_4 = x_{m+1}K$, $P_5 = x_{m+2}K$ and $P_6 = x_{m+3}K \oplus \cdots \oplus x_nK$. As $m \geqq 3$, all subspaces P_i are not zero. Let

$$V^+ = P_1 \oplus P_2 \oplus P_6, \quad V^- = P_4 \oplus P_5 \oplus P_3,$$
$$W^+ = P_1 \oplus P_5 \oplus P_3, \quad W^- = P_2 \oplus P_4 \oplus P_6.$$

Then V^+, V^-, W^+, and W^- are all totally isotropic since the summands P_i in the decomposition of these spaces are all orthogonal. Let v and w be the projective involutions of the second kind in $U_n(K, f)$ whose eigenspaces are $V^\pm$ and $W^\pm$, respectively, and which are conjugate to u. Then u, v, and w all commute and $u^2 = v^2 = w^2 = \beta$. Thus $u_0 = uvw$ is a projective involution of the second kind as ${u_0}^2 = \beta^3$. Hence ${u_0}^*$, v^*, and w^* are conjugate to u^*. Since u^*, v^*, and w^* commute, Theorem 5.2 is proved.

5.3. Involutions of the third kind. Let now u^* be an involution of the third kind in $PU_n(K, f)$ and assume that $n > 4$. We shall show that $\omega_{u^*} > 8$. Let u be a representative of u^*. Then $u^2 = \gamma$ with γ in C and $\gamma = \lambda^2$, where λ is not in Z. As in Section 1.8, let K_1 be the fixed subsfield of K by λ^*. Then $K = K_1 \oplus K_1\theta$, where θ is an element of K which anti-commutes with λ. Then E can be considered as a K_1-space E^*, and $E^* = U^+ \oplus U^-$, where U^+ and U^- are the positive and negative eigenspaces of u. Set $E_1 = U^+$ and form the group $\Gamma U_n(K_1, f_1)$ acting on E_1.

If $\mathfrak{S}^*$ is a set of involutions in $PU_n(K_1, f_1)$, let $c^*(\mathfrak{S}^*)$ denote those involutions of $PU_n(K_1, f_1)$ which commute with every involution of $\mathfrak{S}^*$. If ${v_1}^*$ and ${w_1}^*$ are distinct commuting involutions in $PU_n(K_1, f_1)$, denote the order of $c^*(c^*({v_1}^*, {w_1}^*))$ by $\nu^*({v_1}^*, {w_1}^*)$.

By Theorem 2.2, Section 2.3, there exist two involutions of the first kind in $PU_n(K_1, f_1)$, which we here denote by ${v_1}^*$ and ${w_1}^*$ such that $\nu({v_1}^*, {w_1}^*) > 4$. So we can find an involution ${r_1}^*$ of the first kind (which is one of the ${r_i}^*$ of Section 2.3) in $c^*(c^*({v_1}^*, {w_1}^*))$ which is not one of 1^*, ${v_1}^*$, ${w_1}^*$ or ${v_1}^*{w_1}^*$. The restriction of u to E_1 is a homothetic mapping h_λ in the center ${C_n}^*$ of $U_n(K_1, f_1)$. Let ${v_1}'$, ${w_1}'$ and ${r_1}'$ be the p-involutions ($p \leqq [n/2]$) contained in ${v_1}^*$, ${w_1}^*$, and ${r_1}^*$, respectively; set $v_1 = h_\lambda {v_1}'$ and $w_1 = h_\lambda {w_1}'$. We still have v_1 in ${v_1}^*$ and w_1 in ${w_1}^*$; so r_1 is in $c^*(c^*({v_1}^*, {w_1}^*))$.

We can extend r_1 to an involution r of $U_n(K, f)$ by setting $r(x\theta) = r_1(x)\theta$ for x in $E_1 = U^+$. This defines $r(x)$ on $U^- = U^+\theta$ and consequently on E as $E^* = U^+ \oplus U^-$. Similarly, we can extend v_1 and w_1 to projective involutions

v and w in $U_n(K,f)$. But now $v^2 = w^2 = \gamma$; so v and w are projective involutions of the third kind in $U_n(K,f)$. Therefore, by Theorem 3.3, v and w are in $\mathfrak{D}$.

Since h_λ, v_1 and w_1 all commute, their extensions u, v, and w must all commute. So form $c(c(u,v,w))$. We shall show that r is in $c(c(u,v,w))$. It will then follow that $\omega(u^*,v^*,w^*) > 8$, as the restriction r_1 of r to E_1 does not coincide with the restriction of any of the elements of 1^*, u^*, v^*, w^*, u^*v^*, v^*w^*, w^*u^*, and $u^*v^*w^*$, which are in the cosets 1^*, v_1^*, w_1^* and $v_1^*w_1^*$ of $PU_n(K_1,f_1)$.

If t is a projective involution in $c(u,v,w)$ which commutes with u, then the restriction t_1 of t to E_1 will be a projective involution in $c(v_1,w_1)$. By Section 2.3, $t_1r_1 = r_1t_1$. Let x be in U^+ so that $x\theta$ is in U^-. Then $t(r(x\theta)) = t_1(r_1(x))\theta = r_1(t_1(x))\theta = r(t(x\theta))$. Since $E = U^+ \oplus U^-$, t commutes with r. If t is a projective involution in $c(u,v,w)$ which anticommutes with u, then set $t' = h_\theta t$. Thus t' will commute with u and commute or anticommute with v and w. The restriction t_1 of t' to E_1 will then commute or anticommute with v_1 and w_1. But the subspaces P_i were chosen in Section 2.3 so that no transformation could anticommute with v_1 or w_1. So t_1 commutes with v_1 and w_1 and hence leaves the eigenspaces of r_1 invariant. Thus $t_1r_1 = r_1t_1$. Again let x be in U^+ so that $x\theta$ is in U^-. Then $t'(r(x\theta)) = t_1(r_1(x))\theta = r_1(t_1(x))\theta = r(t'(x\theta))$. This means that t' commutes with r and so t commutes with r. Thus r is in $c(c(u,v,w))$ and $\omega(u^*,v^*,w^*) > 8$.

THEOREM 5.3. *If the dimension of E is greater than 4 and if u^* is an involution of the third kind in $PU_n(K,f)$, $\omega_{u^*} > 8$.*

5.4. Involutions of the fourth kind. We assume that $n > 4$ and E is not commensurable when $n = 6$. We can show that $\mathcal{J}^*$ contains no involutions of the fourth kind when $n > 12$. However, the following argument yields a more general result.

Theorems 5.1, 5.2, and 5.3 imply that $\mathcal{J}^* \cap \mathcal{K}^*$ contains only the extremal involutions and possibly the involutions of the fourth kind. We denote by $\mathfrak{M}(u^*)$ a maximal set of mutually commuting involutions of $\mathcal{J}^* \cap \mathcal{K}^*$ containing u^*. By Lemma 2.1, every involution of $\mathfrak{M}(u^*)$ is extremal or none is. By Section 1.6, every maximal set $\mathfrak{M}(u^*)$ of extremal involutions contains n elements. Furthermore the product of any two or three such involutions is a 2- or 3-involution. Consequently these products are not in $\mathfrak{D}^*$ by virtue of Theorem 2.2. We therefore define $\mathcal{L}^*$ to be the subset of $\mathcal{J}^* \cap \mathcal{K}^*$ consisting of the elements u^* such that the sets $\mathfrak{M}(u^*)$

always possess n elements and such that the product of any two or three involutions in a set $\mathfrak{M}(u^*)$ is not in $\mathfrak{J}^*$. Our principal result is the following theorem.

THEOREM 5.4. *If the dimension n of E is greater than 4 and E is not commensurable when $n = 6$, then the set $\mathfrak{L}^*$ is the set of extremal involutions.*

In view of the preceding remarks, it shall suffice to show that if u^* is an involution of the fourth kind and $\mathfrak{M}(u^*)$ contains n elements, then there exists a product of 2 or 3 involutions of $\mathfrak{M}(u^*)$ in $\mathfrak{J}^*$. Let $u^* = u_1^*, u_2^*, \cdots, u_n^*$ be the elements of $\mathfrak{M}(u^*)$ and u_i be representatives for u_i^*, $i = 1, 2, \cdots, n$. Then $u_iu_j = \pm u_ju_i$, and $u_i^2 = \gamma_i$, where γ_i is not a square in K. Set $(u_iu_j)^2 = \gamma_{ij} = \pm \gamma_i\gamma_j$. First consider that there is a pair i and j, $i \neq j$, such that γ_{ij} is not a square in Z. Then u_iu_j is a projective involution of the third or fourth kinds and is consequently in $\mathfrak{J}$ by Theorems 3.1 and 4.1.

On the other hand, suppose $\gamma_{ij} = \lambda_{ij}^2$ with λ_{ij} in Z for all i and j, where $i \neq j$. If then there exists a projective involution u_k, $k > 2$, which commutes with u_1u_2, then $u_1u_2u_k$ is of the fourth kind as $\lambda_{12}^2\gamma_k$ cannot be a square in K. Finally suppose that all u_k, $k > 2$, anticommute with u_1u_2. Since $n > 4$, there exist two u_k, say u_3 and u_4 which commute with one, say u_1, and anticommute with the other. But then u_1 commutes with u_3u_4, and by the previous argument $u_1u_3u_4$ is in $\mathfrak{J}$. This proves the theorem.

5.5. E commensurable and $n = 6$. Designate by π_{u^*} the maximum of the numbers of involutions in the sets $\mathfrak{M}(u^*)$ of mutually commuting and conjugate involutions containing u^*.

THEOREM 5.5. *If E is commensurable and $n = 6$, then $\pi_{u^*} = 6$ for extremal involutions u^* in $PU_n(K, f)$ and $\pi_{u^*} > 6$ for nonextremal involutions u^* of the first kind.*

Because E is commensurable, all extremal involutions are conjugate to each other. By Section 1.6, the maximum number of such involutions in a set of mutually commuting involutions is 6. Therefore $\pi_{u^*} = 6$ if u^* is extremal.

Similarly, all 2-involutions [3-involutions] of $PU_n(K, f)$ are conjugate. Dieudonné [5; p. 5] has shown that there exist 15 2-involutions [20 3-involutions] in a maximal set of commuting 2-involutions [3-involutions] of $GL_n(K)$. The same result is true for $U_n(K, f)$ by the same argument. These 2-involutions [3-involutions] come from 15 [10] distinct cosets as there is but one [two] involutions in a coset which is a 2-involution [3-involution] of $PU_n(K, f)$. Thus $\pi_{u^*} > 6$ when u^* is a 2-involution [3-involution].

THEOREM 5.6. *When the dimension n of E is 6 and E is commensurable, then the set of extremal involutions in $PU_n(K,f)$ can be characterized as the set of involutions u^* for which*

(a) $\pi_{u^*} = 6$.

(b) *The product of u^* with any two other involutions from any set $\mathfrak{N}(u^*)$ of mutually commuting involutions conjugate to u^* is never conjugate to u^*.*

(c) *Any product of k distinct involutions from $\mathfrak{N}(u^*)$ is conjugate to any other product of k distinct involutions from $\mathfrak{N}(u^*)$.*

(d) *The product of all the involutions in $\mathfrak{N}(u^*)$ is 1^*.*

(e) *For any three involutions v_1^*, v_2^*, and v_3^* in $\mathfrak{N}(u^*)$, $\omega(v_1^*, v_2^*, v_3^*) = 8$.*

We first mention that extremal involutions u^* satisfy the above conditions. Indeed, Theorem 5.5 guarantees (a) is satisfied. Theorem 5.2 justifies (b). From Section 1.6, we see that any product of k distinct extremal involutions is a p- or $(n-p)$-involution, and from Section 1.5 all such involutions are conjugate. The justification for (d) is in Section 1.6. Finally, in the proof of Theorem 5.1, we established (e); the assumption that E be not commensurable was used there only to exclude 3-involutions from the set $\mathcal{J}^*$ from which we chose the involutions v^* and w^* commuting with u^*.

Theorem 5.5 states that the 2- and 3-involutions of $PU_6(K,f)$ do not satisfy (a). Theorem 5.2 states that the involutions of the second kind in $PU_6(K,f)$ do not satisfy (b). So we must only consider the cases where u^* is an involution of the third or fourth kind. We show that not all the conditions of Theorem 5.6 are satisfied.

We assume that u^* satisfies (a), (b), (c), and (d). Thus since $\pi_{u^*} = 6$, we can pick a set $\mathfrak{N}(u^*)$ of 6 mutually commuting involutions conjugate to u^*, one of which is u^*. Let $u_1^* = u^*, u_2^*, \cdots, u_6^*$ be the elements of $\mathfrak{N}(u^*)$. Pick representatives u_i in u_i^*. By (d), we have that $u_1u_2u_3u_4u_5u_6 = h_\lambda$ in C_n. This means that u_1 commutes with $u_2u_3u_4u_5u_6$ and thus commutes with one u_i, $i > 1$, say u_2. It then follows that $w_1 = u_1u_2$ is a projective involution of the first kind. Consequently by (c), $w_2 = u_3u_4$ and $w_3 = u_5u_6$ are in cosets w_2^* and w_3^*, respectively, which are conjugate to w_1^* when considered as elements of $PU_6(K,f)$. Since $w_1^*w_2^*w_3^* = 1^*$, it is possible to choose the representatives w_i and u_i so that w_i, $i = 1, 2, 3$, are involutions of $U_n(K,f)$ and $w_1w_2w_3 = 1$.

We shall show that the involutions w_i are 2-involutions. Indeed, as they come from conjugate cosets, they are all either 1-involutions, 2-involutions or 3-involutions. However, we have that $w_1 = w_2 w_3$.

In the case that the involutions w_i are 1-involutions or 3-involutions, $w_2 w_3$ is easily verified to be a 2-involution while $w_1 = w_2 w_3$. So it follows that w_i are all 2-involutions. Then by (c) $u_i^* u_j^*$ is a 2-involution for any two distinct involutions u_i^*, u_j^*. Thus $u_i u_j$ is a projective involution of the first kind in $U_n(K, f)$ which is in the coset of a 2-involution; that is, $u_i u_j$ have a 2-dimension eigenspace. As $n = 6$, no element t in $U_6(K, f)$ anti-commutes with $u_i u_j$. This means that all the projective involutions u_i commute; for, should $u_i u_j = -u_j u_i$, then would $u_i(u_i u_j) = -(u_i u_j) u_i$.

Let now u^* be an involution of the third kind, and form the group $\Gamma U_6(K_1, f_1)$ acting on the K_1-subspace of E defined in Section 1.8. Since $u_i u_1$, $i > 1$, is in the coset of a 2-involution of $PU_6(K, f)$, u_i is represented in $U_6(K_1, f_1)$ by a projective involution u_i' of the first kind with a 2-dimensional eigenspace U_i^+ of E_1 and 4-dimensional eigenspace U_i^-. As there are five such involutions u_i', we can find two, say u_2 and u_3, such that $U_2^+ \cap U_3^+$ is 1-dimensional. It then follows that none of the spaces $U_2^\pm \cap U_3^\pm$ are zero. Hence by the argument of Section 5.3, $\omega(u_1^*, u_2^*, u_3) > 8$. Thus u^* does not fulfill condition (e).

Now assume that u^* is of the fourth kind in $PU_6(K, f)$. As in Section 1.9, form the group $\Gamma U_3(L, g)$ acting on the L-vector space F. In $U_6(K, f)$ the projective involutions $u_i u_1$ have distinct 2- and 4-dimensional eigenspaces. Considered as elements u_i' of $U_3(L, g)$, they are 1-involutions as the L-dimension of a subspace is half its K-dimension. Furthermore, u_i' and u_j' always represent distinct cosets u_i^* and u_j^* from $PU_3(L, g)$; for, otherwise, their eigenspaces would coincide as subspaces of F. This would mean that the eigenspaces of $u_i u_1$ and $u_j u_1$ would coincide as subspaces of E. Then we would have $u_i^* u_1^* = u_j^* u_1^*$; that is, $u_i^* u_j^* = 1^*$, which is contradiction to (c).

This implies that $u'_2{}^*, u'_3{}^*, \cdots, u'_6{}^*$ form a set of distinct mutually commuting extremal involutions in $PU_3(L, g)$. Since the maximum number of such involutions is 3 by Section 1.6, we have a contradiction. This establishes Theorem 5.6.

UNIVERSITY OF MICHIGAN AND UNIVERSITY OF WASHINGTON.

REFERENCES.

[1] G. Birkhoff and J. von Neumann, "The logic of quantum mechanics," *Annals of Mathematics*, vol. 37 (1936), pp. 823-243.

[2] J. Dieudonné, "La géometrie des groupes classiques," *Ergebnisse der Mathematik und ihrer Grenzgebeite*, Neue Folge, Heft 5 (1955), Berlin (Springer-Verlag).

[3] ———, "Les extension quadratiques des corps non commutatifs et leurs applications," *Acta Mathematica*, vol. 87 (1952), pp. 175-242.

[4] ———, "Sur les groupes classiques," *Actualités Scientifiques et Industrielles*, No. 1040 (1950), Paris (Hermann et Cie).

[5] ———, "On the automorphisms of the classical groups," *Memoirs of the American Mathematical Society*, No. 2 (1950).

[6] G. W. Mackey, "Isomorphisms of normed linear spaces," *Annals of Mathematics*, vol. 43 (1942), pp. 244-260.

[7] C. E. Rickart, "Isomorphic groups of linear transformations," *American Journal of Mathematics*, vol. 72 (1950), pp. 451-464.

[8] ———, "Isomorphic groups of linear transformations II," *American Journal of Mathematics*, vol. 73 (1951), pp. 697-716.

[9] ———, "Isomorphisms of infinite-dimensional analogues of the classical groups," *Bulletin of the American Mathematical Society*, vol. 57, No. 6 (1951), pp. 435-448.

ON THE LOCAL EMBEDDING PROBLEMS IN THE DIFFERENTIAL GEOMETRY OF SURFACES.*

By AUREL WINTNER.

Let D be an open, simply connected domain in a (u^1, u^2)-plane. For the purposes at hand, it can always be assumed that D is sufficiently small, in the sense of being a sufficiently small neighborhood of a given point, say of $(u^1, u^2) = (0, 0)$. If the elements $g_{ik} = g_{ik}(u^1, u^2)$ of a binary, symmetric matrix (g_{ik}) of non-vanishing determinant are given on D as functions of class C^2, then

$$g_{ik}(u^1, u^2)\,du^i du^k \tag{1}$$

will be called a *C^2-metric* (on D). Then, since $g^{ik} = (g_{ik})^{-1}$ and the first and second derivatives of $g_{ik}(u^1, u^2)$ exist and are continuous, (1) has a continuous curvature $K = K(u^1, u^2)$. Suppose that, for some reason, the curvature K turns out to be of class C^1 (rather than just continuous) on D. Then the C^2-metric (1) will be called a *regular C^2-metric* (on D).

It was shown in [12] that, whereas certain theorems in differential geometry are false for C^2-metrics, they become true for regular C^2-metrics. The purpose of this note is to show that the notion of a regular C^2-metric (or of the corresponding notion of a regular C^n-metric, where $n \geqq 1$) leads to improved formulations of certain solved, and to corresponding sharp formulations of certain unsolved, problems in the theory of local embeddings.

It will be assumed throughout that the metric (1), instead of being just of non-vanishing determinant, is *positive definite*, that is,

$$\det g_{ik} > 0 \text{ and } g_{ii} > 0. \tag{2}$$

Let S be a surface of class C^3 in an (x, y, z)-space and let S be a "small," simply connected, open piece of a surface ("small" in the same sense in which the domain D was defined to be "small"). Then S has a one-to-one parametrization,

$$S\colon\ X = X(u^1, u^2), \qquad X = (x, y, z), \tag{3}$$

which is a C^3-parametrization (so that (u^1, u^2) varies over a D on which

* Received July 27, 1955.

the vector function X is of class C^3 and has, with respect to u^1 and u^2, a (3 by 2) Jacobian matrix of rank 2 or, what is the same thing, the first fundamental form

$$|\,dX(u^1,u^2)\,|^2 \equiv g_{ik}(u^1,u^2)\,du^i du^k \tag{4}$$

is of non-vanishing determinant on D). Only such parametrizations (3) of the C^3-surface will be used as satisfy these conditions. Corresponding definitions and conventions will be used in connection with C^n-surfaces S for any n.

The considerations will center about the following fact:

(i) *If* (3) *is a* C^3*-surface, then its first fundamental form* (4), *instead of being just a* C^2*-metric, is a regular* C^2*-metric.*

The truth of this fact (i) becomes clear if the *theorema egregium* is read in the direction opposite to that emphasized by Gauss (cf. [12], p. 686).

If a metric (1), rather than a surface S, is given, then by an *embedding of* (1) is meant any surface (3) satisfying (4) (provided that there exists such a surface). Clearly, (i) can be interpreted as follows: In order that a metric admits some C^3-embedding, it is necessary that the metric be a regular C^2-metric. It is natural to ask whether this necessary condition is sufficient as well, at least if the given metric has no parabolic points, that is, if the curvature $K(u^1,u^2)$ of the given regular C^2-metric is either positive throughout or negative throughout. It turns out that this converse of (i) is false, even if

$$K(u^1,u^2) \neq 0 \text{ on } D. \tag{5}$$

The failure of the converse of (i) would be less surprising if it were not claimed even under the assumption (5). In fact, if $K(u^1,u^2)$ changes sign on D, then the Monge-Ampère equation, on which the problem of embedding depends, is in a state of transition from the elliptic ($K>0$) to the hyperbolic ($K<0$) type of a partial differential equation of second order, and there are no general existence theorems in such a situation of the Tricomi type (except in the analytic case of Cauchy-Kowalewski). A corresponding remark holds if the metric (1), given on D, fails to satisfy (5) but is either non-hyperbolic ($K \geqq 0$) or non-elliptic ($K \leqq 0$), without being parabolic ($K \equiv 0$).

What *is* true in the case (5) is the following *partial* converse of (i):

(ii) *Every regular* C^2*-metric* (1) *which is of non-vanishing curvature* (*that is, which is either hyperbolic,*

$$K(u^1,u^2) < 0, \tag{6}$$

or elliptic,

$$K(u^1, u^2) > 0, \tag{7}$$

on D) *possesses* C^2*-embeddings* S (*locally*; that is, embeddings of D_0, if D_0 is a sufficiently small neighborhood of any given point of D).

What *is not* true is a certain *full* converse (i) in the case (5), since the situation as follows:

(iii) *Under the assumptions of* (ii) *on the metric* (1), *a* C^2*-embedding* S, *the existence of which is assured by* (ii), *can fail to be a* C^3*-embedding not only in the hyperbolic case but also in the elliptic case.*

It will remain *undecided* whether a *complete* converse of (i) in the case (5) is false in the following sense also:

(iii ?) *Under the assumptions of* (ii) *on the metric* (1), *must there exist some* C^3*-embedding* S *in the hyperbolic case and/or in the elliptic case?*

Actually, the answer to the questions (iii ?) does not seem to be known even if (1) is not restricted by the assumption (5). But if (5) is omitted, then nothing is known, except in the analytic case. In fact, even the following question (representing the weakest form of anything like (ii) when (5) is not assumed) is unanswered:

(ii ?) *Does every* C^∞*-metric possess some* C^∞*-embedding or, at least, some* C^2*-embedding?*

It is understood that (1) is called a C^∞-metric if the three functions $g_{ik}(u^1, u^2)$ possess continuous partial derivatives of arbitrarily high order (then the same is true of $K(u^1, u^2)$; so that every "C^∞-metric" is a "regular C^∞-metric"). The trouble is that there is not available an existence theorem (not even with suitable, rather than with arbitrary, Cauchy data) for a system of partial differential equations which has the same form as a Cauchy-Kowalewski system but in which the coefficient functions, instead of being analytic, are just of class C^∞. As a matter of fact, such an existence theorem seems to be doubtful even in case of *linear* Cauchy-Kowalewski systems with coefficient functions of class C^∞.

In the analytic case, the answer to the analogue of question (ii ?) is known. In fact, every analytic metric (1) possesses analytic embeddings. This follows from the Cauchy-Kowalewski theorem, since the latter can be applied to the problem of embedding so as to avoid any reference to assumption (5).

In the *elliptic* case of question (iii ?), there is an affirmative answer

under a "restriction of comparison." This is the content of the following criterion (iv):

(iv) *If an elliptic, regular C^2-metric possesses a C^1-embedding, then all of its C^2-embeddings (which, by* (ii)*, exist) are C^3-embeddings.*

It is clear from the case $K \equiv -1$ of an analytic metric that (iv) becomes false if its assumption (7) is replaced by (6). The following question will remain unanswered:

(iv bis ?) *Does* (iv) *remain true if its C^1 is reduced to C^3 ?*

In (iv ?) below, the regularity of a C^1-metric is meant in the sense of the existence of a continuous curvature $K = K(u^1, v^2)$ for the C^1-metric (1), when K is defined in the sense of Weyl [9], pp. 42-44; cf. the comments preceding (8) below.

(iv) suggests a question concerning an elliptic analogue of (ii). But this question remains undecided:

(iv ?) *Must every elliptic, regular C^1-metric possess an embedding which is a surface S of class C^2 ?*

A positive answer to this question (iv ?) would be important, since if the answer were in the negative, it would follow that, *for local reasons*, Weyl's embedding problem *in the large* cannot be solved with that generality with which he has stated it originally (cf. [6], p. 485 and p. 487, and the footnote in [12], p. 685).

Corresponding to the two cases, (6) and (7), of (ii), there is a hyperbolic counterpart of (iv bis ?). Since no potential-theoretical re-enforcements can be forthcoming in this hyperbolic case, it appears that the answer to this counterpart of (iv bis ?) is in the negative (a counterexample, if any, seems to be harder in the elliptic case).

The following pair of problems is of quite another character.

(+ ?), (— ?) *Let S be a C^1-embedding of a metric* (1) *on D which is an analytic metric or, for that matter, just a regular C^1-metric* (*so that* (1) *defines a continuous curvature*). *Must then S be a convex surface or a saddle surface over D according as the curvature of* (1) *satisfies* (7) *or* (6)?

It is understood that convexity and saddle character of C^1-surfaces are defined in terms of the tangent plane (which, locally, supports S in the first case but not in the second). If S is a C^2-surface, then it has a (continuous) second fundamental form, and so the questions are trivial. But if S is of

class C^1 without being of class C^2, then (7) or (6) cannot be expressed in terms of a second functional form, since the latter fails to exist.

It will be clear from the proofs that both (ii) and (iii) remain true if C^2 and C^3 are replaced by C^n and C^{n+1}, respectively, where $n \geqq 2$. A corresponding remark holds, of course, in (i), where $n \geqq 1$, whereas the truth of the analogues of (ii) and (iii) will remain undecided if $n = 1$. It is understood that by a regular C^n-metric is meant a C^n-metric possessing a curvature $K(u^1, u^2)$ which is a function of class C^{n-1}.

The general concept and the name of a regular C^1-metric are due to H. Weyl and H. Cartan, respectively; for references, cf. [2]. For metrics (1) which are given in an isothermic form (that is, in a form in which

$$g_{11} = g_{22} \text{ and } g_{12} = 0 \tag{8}$$

hold), the concept of a regular C^1-metric is contained already in a paper of Lichtenstein [8], pp. 87-88 and 96-100 (in fact, cf. the relation to which formula (15) in [12] reduces in the isothermic case). This is the more interesting since, as it turned out in the meantime, there is no loss of generality in assuming (8). In fact, it was shown in [2] that, after a suitable (u^1, u^2)-transformation (which is of class C^2 and of non-vanishing Jacobian), *every* regular C^1-metric (1) becomes a regular C^1-metric satisfying (8).

Proof of (ii). Consider Darboux's differential equation (cf. [1], pp. 240-241) for an unknown function $z = z(u, v)$ of class C^2 of the embedding problem, where, if (1) is the given metric (to be embedded),

$$u = u^1, \qquad v = u^2. \tag{9}$$

Let $p, \cdots, t$ denote the partial derivatives $z_u, \cdots, z_{vv}$. Darboux's differential equation for $z = z(u, v)$ is a Monge-Ampère equation,

$$g^2 \cdot (rt - s^2) + ar + 2bs + ct + d = 0, \tag{10}$$

in which the coefficients $g^2 = \det g_{ik} > 0$, a, b, c, d do not contain z, are polynomials in p, q, contain the independent variables u, v only in terms of the (given) functions

$$g_{ik}(u, v), g_{ik\,u}(u, v), g_{ik\,v}(u, v), K(u, v), \tag{11}$$

and are polynomials in the latter. It is also well-known that the bicharacteristic type of (10) is hyperbolic or elliptic according as the metric (1) is hyperbolic or elliptic. In view of (5), this means that the partial differential equation (10) is of hyperbolic or of elliptic type according as

$$K(u, v) < 0, \tag{12-}$$

$$K(u, v) > 0. \tag{12+}$$

Since (ii) assumes that (1) is a regular C^2-metric, all functions (11) are of class C^1 in (u, v). Hence, it is possible to apply to (10) Theorem (V), p. 849, of [4] on hyperbolic Monge-Ampère equations in the case (12—) and Theorem (§), pp. 316-317, of [5] on general (not necessarily Monge-Ampère) elliptic equations in the case (12+) (in which case, instead of the C^1-character of the functions (11), their C^λ-character, where λ is a Hölder index satisfying $0 < \lambda < 1$, would suffice; cf. (§), *loc. cit.*). Thus, the pair of existence theorems, just quoted, leads to the following result:

If (1) is a regular C^2-metric on a (u, v)-domain D, where $(u, v) = (u^1, u^2)$, then, corresponding to every non-parabolic point (u_0, v_0) of D, there exists a neighborhood D_0 of (u_0, v_0) on which Darboux's equation (10) posesses solutions $z = z(u, v)$ of class C^2. In addition, since such a C^2-solution $z = z(u, v)$ belongs to every sufficiently smooth choice of Cauchy initial data in the hyperbolic case and of Dirichlet boundary data in the elliptic case (cf. *loc. cit.*), these initial or boundary data can be chosen so "small" that, if the corresponding C^2-solution $z(u, v)$ is subjected to the first Beltrami operator ∇ of the given metric (1), then the resulting function $\nabla z(u, v)$ (which clearly is a function of class C^1 on D_0 and, since (1) is positive definite, attains non-negative values only) will satisfy the inequality

$$\nabla z(u, v) < 1 \tag{13}$$

at every point (u, v) of D_0 (if the neighborhood D_0 of (u_0, v_0) is small enough).

The existence of a surface S having the properties claimed by (ii) now follows, by Darboux's procedure (cf. the first italicized statement in [1], p. 241), from the eixstence of functions $z = z(u, v)$ of class C^2 satisfying both (10) and (13) on D_0.

It should be noted that, in view of that device in Darboux's procedure which is expressed by the italicized statement in [1], p. 240, recourse must be had to a certain lemma on "euclidean" mappings. The lemma in question is, however, contained in the statements (I*), (II*) of [12], p. 681.

Remark. Owing to the freedom in choosing the (smooth) Dirichlet data in the case $K > 0$ above, the C^2-solutions $z = z(u, v)$ of (10) can be chosen so as to satisfy, besides (13), other inequalities also (on D_0, if D_0 is small enough). Since such an inequality is $H^2 \neq K$, where $H = H(u, v)$ denotes the mean curvature of a surface S, to be supplied by (ii), and $K = K(u, v)$ is the Gaussian curvature, hence a function given along with (1), it follows that, *under the assumptions of* (ii), *the C^2-embedding S can be chosen so as to be free of umbilical points.*

Proof of (iii). Let (1) be given by $(du)^2 + (e^{cu}dv)^2$, where $(u, v) = (u^1, u^2)$ and $c = \text{const}$. Then the curvature $K(u^1, u^2)$ is $-c$, hence $K \equiv -1$ if $c = 1$. Since $K < 0$, the Monge-Ampère equation (10) belonging to this analytic metric (1) is of hyperbolic type, i. e., it has distinct real bicharacteristics, and possesses therefore solutions $z(u, v)$ which are of class C^n without being of class C^{n+1}, where $n \geqq 2$ can be chosen arbitrarily (for details, cf. a corresponding construction in [5], pp. 329-330).

Thus (iii) is clear in the hyperbolic case ($K < 0$). An example proving (iii) in the elliptic case is less straightforward. Such an example ($K > 0$) was given in [6], p. 486.

Remark. It would be desirable to decide whether such an example ($K > 0$) is possible also when it is required to be *free of umbilical points* (in fact, a perusal of the example given in [6], p. 486, shows that the condition $H^2 \neq K$ is violated at the crucial point of the surface obtained there).

Proof of (iv). The following concepts will be needed:

Two metrics, say (1) and

$$E(u, v)du^2 + 2F(u, v)du\,dv + G(u, v)dv^2, \tag{14}$$

are *C^m-isometric* (for a fixed $m > 0$) if there exists a local transformation T,

$$u = u(u^1, u^2), \qquad v = v(u^1, u^2), \tag{15}$$

in which the functions u, v are of class C^m in (u^1, u^2), have a non-vanishing Jacobian, and are such that (1) and (14) become identical by virtue of (15) (or, equivalently, of its inverse, T^{-1}). Thus, if (14) is a regular C^n-metric for a fixed $n > 0$ (that is, if the coefficients of (14) are functions of class C^n and if, in addition, (14) possesses a curvature $K = K(u, v)$ which is a function of class C^{n-1}), then (14) is C^{n+1}-isometric to a regular C^n-metric (1) satisfying (8). This was proved for $n > 1$ in [12] and for the (more intricate) case $n = 1$ in [2]. The necessity of dealing with the concept of C^m-isometry was pointed out in [11]; cf. also [10]. The answer to the question raised in [11] was answered by Hartman [3] in the affirmative. It is readily seen that Hartman's result contains the following fact:

LEMMA. *Every C^1-isometry T of two regular C^n-metrics is a C^{n+1}-isometry* ($n > 0$ is arbitrarily fixed).

If C^1 and $n > 0$ in this Lemma are reduced to C^2 and $n > 1$, respectively, then a simple proof results from [10]; cf. [12], pp. 681-682. This reduced formulation would suffice in the applications of the Lemma below.

If the C^m-isometry (15) belongs to (1) and to the case

$$E = g_{11}, \qquad F = g_{12}, \qquad G = g_{22} \tag{16}$$

of (14), let T be called a *C^m-autometry* (of (1) or, since T can be replaced by T^{-1}, of (14)). Clearly, the lemma contains the following corollary:

COROLLARY. *Every C^1-autometry T of a regular C^n-metric is a C^{n+1}-autometry* ($n > 0$ is arbitrarily fixed).

Consider two embeddings, say S and S^*, of one and the same metric (1). If one of these two surfaces, say S^*, is of class C^m for a fixed $m > 1$, then its first fundamental form (1) is a regular C^{m-1}-metric. Since S and S^* have the same first fundamental form, it follows from the corollary that the correspondence, induced between S and S^* by their autometry, is represented by a C^m-transformation (15) of non-vanishing Jacobian. Hence (iv) follows from the theorem italicized at the end of p. 612 in [7], if use is made of the Lemma above.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] Bianchi, *Lezioni di geometria differentiale*, 2nd ed. (1902), vol. 1.

[2] S.-S. Chern, P. Hartman and A. Wintner, "On isothermic coordinates," *Commentarii Mathematici Helvetici*, vol. 28 (1954), pp. 301-309.

[3] P. Hartman, "On unsmooth two-dimensional Riemannian metrics," *American Journal of Mathematics*, vol. 74 (1952), pp. 215-226 (Part II, pp. 222-226).

[4] ——— and A. Wintner, "On hyperbolic partial differential equations," *ibid.*, vol. 74 (1952), pp. 834-864.

[5] ——— and A. Wintner, "On the third fundamental form of a surface," *ibid.*, vol. 75 (1953), pp. 298-334.

[6] ——— and A. Wintner, "On pieces of convex surfaces," *ibid.*, vol. 75 (1953), pp. 477-487.

[7] ——— and A. Wintner, "On elliptic Monge-Ampère equations," *ibid.*, vol. 75 (1953), pp. 611-620.

[8] L. Lichtenstein, "Über einige Integrabilitätsbedingungen zweigliedriger Differentialausdrücke mit einer Anwendung auf den Cauchyschen Integralsatz," *Sitzungsberichte der Berliner Mathematischen Gesellschaft*, vol. 9 (1910), pp. 84-100.

[9] H. Weyl, "Über die Bestimmung einer geschlossenen konvexen Fläche durch ihr Linienelement," *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, vol. 61 (1915), pp. 40-72.

[10] A. Wintner, "On Riemann metrics of constant curvature," *American Journal of Mathematics*, vol. 73 (1951), pp. 569-575.

[11] ———, "On isometric surfaces," *ibid.*, vol. 74 (1952), pp. 198-214.

[12] ———, "On the local rôle of the theory of the logarithmic potential in differential geometry," *ibid.*, vol. 75 (1953), pp. 679-690.

ON INDEFINITE BINARY RIEMANNIAN METRICS.*

By AUREL WINTNER.

---

1. The following considerations were suggested by a theorem which, in the nomenclature used in [14], can be formulated as follows:

(i_n) *If a binary, definite C^n-metric*

$$g_{ik}(u^1, u^2)\,du^i du^k \qquad (g_{12} = g_{21}) \tag{1}$$

is a regular C^n-metric, then it is isometric to a C^n-metric having the conformal normal form

$$\lambda(u, v)\,(du^2 + dv^2) \qquad (\lambda \neq 0) \tag{2}$$

(and (2) is, of course, a regular C^n-metric).

Here and in the sequel, all assertions are of a local nature, that is, the (u^1, u^2)- or (u, v)-domain is meant to be sufficiently small (in the sense specified at the beginning of [14]).

(i_n) was proved for $n > 1$ in [13], where the limiting case $n = 1$ was left undecided. It was proved in [2] that (i_n) can be extended to this limiting case (which is more intricate than are the cases $n = 2, 3, \cdots$).

In view of the comments made in [13], pp. 686-687, the case $n = 2$ of (i_n) contains a justification of a tacit assumption made by H. Hopf in connection with his elegant formulation of the Codazzi equations (cf. [6], p. 240, formula (4), and p. 234, line 5).

In order to deal with parametrizations of a surface in terms of its lines of curvature and, when the Gaussian curvature is negative, of its asymptotic lines, there is needed an analogue of (i_n) for the case in which the given metric (1) is indefinite; so that what formally corresponds to the normal form (2) is

$$2\lambda(u, v)\,du dv \qquad (\lambda \neq 0) \tag{3}$$

(in this connection, cf. [11], pp. 204-205). It will be shown in Sections 7-12 that both parametrization problems, just mentioned, can actually be dealt with on the basis of the following theorem:

(ii_n) *If a binary, indefinite C^n-metric* (1) *is a regular C^n-metric, then*

---

* Received July 27, 1955.

it is isometric to a C^n-metric having the "non-euclidean conformal" normal form (3) (*and* (3) *is, of course, a regular C^n-metric*).

(i_n) deals with a problem in partial differential equations of elliptic type. Correspondingly, the proof of (i_n) depended on potential-theoretical arguments. No such analytic reinforcement can be forthcoming for the problem in (ii_n), since the underlying partial differential equation becomes of hyperbolic type. Thus it is quite unexpected that the counterpart (ii_n) of (i_n) turns out to be true.

What is more, the example given in [4] seemed to indicate that in the applications to theory of surfaces, mentioned above, *two* degrees of differentiability must be sacrificed before the normal form of the parametrization can in general be reached. That actually only *one* degree is lost, will turn out to be a consequence of (ii_n) and of the fact that, in the applications in question, the regularity assumption of (ii_n) happens to be satisfied; cf. the Appendix to [13].

Corresponding to the restriction $n > 1$ on (i_n) in [13], the proof of (ii_n) will be given under the restriction $n > 1$. The limiting index $n = 1$, dealt with in the case of (i_n) in [2], leads in the case of (ii_n) to the same technical complications as in the case of (i_n); complications which undoubtedly can be overcome along the lines of considerations carried out in [2] in the elliptic case and in [5], pp. 357-361, in a similar but hyperbolic (Chebyshev parametrizations).

2. The proof of (ii_n) will be given for $n = 2$. The proof will be such as to make clear its validity for every higher n-value.

Let D be a simply connected, open (u^1, u^2)-domain; it can be assumed to be a sufficiently small neighborhood of the point $(u^1, u^2) = (0, 0)$. The assumptions of (ii_2) are that, on D, the coefficients of (1) are functions of class C^2; that $\det g_{ik} < 0$; finally, that the curvature $K = K(u^1, u^2)$ of (1) is a function of class C^1 (rather than just of class $C^{n-2} = C^{2-2}$, i.e., just continuous; in fact, the regularity of the C^2-metric (1) means, by definition, this increase in the degree of differentiability of K).

For the present, only the C^2-character of (1) will be used (so that $K(u^1, u^2)$ can, for the present, be just continuous). Then, since $\det g_{ik}(u^1, u^2) < 0$ on D, it is clear that the condition

$$g_{ik}(u^1, u^2)\, du^i du^k = 0 \tag{4}$$

can be split into a system of the form

$$du^1/du^2 = f(u^1, u^2), \qquad du^2/du^1 = g(u^1, u^2) \tag{5}$$

in which f and g are real-valued functions of class C^2. Hence (5) has a pair of (local) first integrals, say

$$u(u^1, u^2) = \text{const.}, \qquad v(u^1, u^2) = \text{Const.}, \tag{6}$$

where the two functions u, v of (u^1, u^2) can be chosen to be of class C^2 and of non-vanishing Jacobian. Hence, if T denotes the (local) C^2-transformation $(u^1, u^2) \to (u, v)$ which results if const. and Const. in (6) are denoted by u and v, respectively, then T^{-1} exists and is of class C^2. Since (4) is equivalent to the system (5), it is clear from the definition of T that T transforms (1) into the normal form (3), and that

$$\lambda(u, v) \in C^1 \qquad (\lambda \neq 0) \tag{7}$$

in (3); in fact, the transformation T (along with its inverse T^{-1}) and the mertic (1) are of class C^2. The parenthetical assertion of (7) follows, of course, from $\det g_{ik} \neq 0$.

By using the assumption that (1) is a *regular* C^2-metric, an assumption which was not used thus far, it will be shown in Sections 3-5 that (7) can be improved to

$$\lambda(u, v) \in C^2 \qquad (\lambda \neq 0), \tag{8}$$

provided that the C^2-transformation T is suitably chosen. Since (8) means that (3) is a C^2-metric, this will prove (ii_2). In fact, the parenthetical assertion made after (3) is clear from the invariance of the scalar $K(u^1, u^2) \in C^1$ under the mapping T.

3. According to the result (7) of Section 2, the regular C^2-metric (1) is C^2-isometric to a regular C^1-metric (3). If $K(u, v)$ denotes the curvature of (3), then

$$K(u, v) \in C^1 \tag{9}$$

(by invariance; cf. the end of Section 2). But since (8) is not known, the classical representation of the curvature of (3), the representation supplied by the formula

$$(\log \lambda)_{uv} = -K\lambda, \tag{10}$$

where $_z = \partial\ /\partial z$ (cf., e.g., [1], formula (IV) on p. 52), cannot be applied. What can be applied to the C^1-parametrization (3) of the C^2-metric (1) is the integrated form of (10), that is, Weyl's definition ([10], pp. 42-44) of the curvature of a regular C^1-metric; cf., e.g., formula (15) in [13], p. 682 (cf. also the theorem (II) in [5], p. 348, which deals with a more general

situation). In the case of the indefinite metric (3), that "integrated form" of the definition of K reduces to

$$\int_J \{(\log\lambda)_v\,dv - (\log\lambda)_u\,du\} = \iint_B -2K\lambda\,dudv \tag{11}$$

(cf. (10) or the formula just quoted). Here J denotes any positively oriented, piecewise smooth Jordan curve which, along with its interior domain $B = B(J)$ which occurs on the right of (11), is contained in the (u, v)-domain under consideration.

Choose J to be the boundary of the rectangle B having two points (u, v), $(u + \Delta u, v + \Delta v)$ as opposite vertices. Then the line integral in (11) is readily seen to reduce to $2\square\log\lambda$, if $\square\phi$ is defined by

$$\square\phi = \phi(u + \Delta u, v + \Delta v) - \phi(u, v + \Delta v) - \phi(u + \Delta u, v) + \phi(u, v). \tag{12}$$

Hence (11) is equivalent to the following identity in (u, v) and $(\Delta u, \Delta v)$:

$$\square\log\lambda = \int_u^{u+\Delta u}\int_v^{v+\Delta v} -K\lambda\,dudv. \tag{13}$$

Since K and λ are continuous, (13) implies that, as $(\Delta u, \Delta v) \to (0, 0)$, the limit relation

$$(\square\log\lambda)/(\Delta u\cdot\Delta v) \to -K(u, v)\lambda(u, v) \tag{14}$$

holds at every point (u, v) of the (u, v)-domain under consideration.

Actually, it is clear from (7), (9), (13) and (12) that $(\log\lambda)_{uv}$ exists and satisfies (10) (so that, in particular, $\lambda_{uv} = \lambda_{vu}$ exists and is continuous). What remains to be shown is that, if the (u, v)-parametrization is suitably chosen, then (10) and (9) improve (7) to (8).

4. To this end, define a function γ of (u, v) by placing

$$\gamma(u, v) = \int^u\int^v -K\lambda\,dudv, \tag{15}$$

where the choice of the lower limits of integration will be immaterial. It is seen from (7), (9) and (15) that

$$\gamma(u, v)\ \varepsilon\ C^2. \tag{16}$$

On the other hand, it is clear from (15) that $\gamma_{uv} = -K\lambda$. It follows therefore from (10) that the difference

$$\phi = \log\lambda - \gamma \tag{17}$$

satisfies the relation

$$\phi_{uv} = 0. \tag{18}$$

Since (18) means that $\phi = \phi(u, v)$ is of the form $\alpha(u) + \beta(v)$, it follows from (17) that

$$\log \lambda(u, v) = \gamma(u, v) + \alpha(u) + \beta(v). \tag{19}$$

Hence, from (7) and (16),

$$\alpha(u) \in C^1, \qquad \beta(v) \in C^1. \tag{20}$$

If (19) and (20) are compared with (7) and (16), it follows again (cf. the end of Section 3) that

$$\lambda_{uv}(u, v) = \lambda_{vu}(u, v) \text{ exists and is continuous.} \tag{21}$$

5. Since (10) is a partial differential equation of hyperbolic type, (7) and (21) together contain much less than (8), the assertion to be proved. But it will now be shown that the goal (8) can be reached by an adaptation of a device repeatedly used in [12].

Let the notation be changed in such a way that $\log \gamma$, $\log \alpha$, $\log \beta$ become the functions which in (19) are denoted by γ, α, β, respectively. Then (16) and (20) hold for the new functions also, the new functions are positive, and (3) goes over into

$$2\gamma(u, v)\alpha(u)\beta(v)\,du\,dv, \tag{22}$$

since (19) becomes $\lambda = \gamma\alpha\beta$.

Next, since $\alpha(u)$, $\beta(v)$ are positive and satisfy (20), it is clear that

$$u' = \int^u \alpha(u)\,du, \qquad v' = \int^v \beta(v)\,dv \tag{23}$$

defines a C^2-transformation $(u, v) \to (u', v')$ of non-vanishing Jacobian. Hence, if a function λ' of (u', v') be defined by placing $\lambda'(u', v') = \gamma(u, v)$ (by virtue of (23)), then (16) implies that

$$\lambda'(u', v') \in C^2, \tag{24}$$

whilst (23) shows that (22) is identical with

$$2\lambda'(u', v')\,du'\,dv' \qquad (\lambda' > 0). \tag{25}$$

Finally, if the primes are omitted, then (25) and (24) go over into (3) and (8), respectively. The proof is therefore complete.

6. Applications of (ii_n) to the classical parametrization problems of surfaces

$$S\colon X = X(u,v), \text{ where } X = (x,y,z), \tag{26}$$

will now be given. The following terminology will be used:

By a surface S will be meant a sufficiently small piece of a surface, that is, some neighborhood of any given point of the "surface." If S has a parametrization (26) in which the vector $X(u,v)$ is, for some $m \geqq 1$, a function of class C^m on a (sufficiently small) (u,v)-domain D, and if the vector product of the partial derivatives X_u, X_v does not vanish on D, then (26) will be called a C^m-parametrization of S, and S will be referred to as a surface of class C^m (in symbols: $S \in C^m$).

Note that $k < m$ is quite possible if (26) is a C^k-parametrization of an $S \in C^m$ (even if the best possible value of k is chosen for the *given* parametrization (21) of $S \in C^m$). It is precisely this possibility which creates the issues treated below.

7. The following theorem will first be proved:

(iii_m) *If the Gaussian curvature is negative on an* $S \in C^m$, *then* S *possesses* C^{m-1}*-parametrizations* (26) *in which the parameter lines*

$$u = \text{const.}, \qquad v = \text{Const.} \tag{27}$$

are asymptotic lines of S.

The proof of (iii_m) will depend on the case $n = m-2$ of (ii_n). Hence (iii_m) will follow for $m > 3$, since (ii_n) was proved for $n > 1$. Correspondingly, (iii_3) is true if (i_1) is. On the other hand, (ii_2) is false (in every sense); cf. [4]. This failure will be seen to be due to the fact that nothing like the result of the Appendix of [13] on an $S \in C^m$, where $m > 2$, can even be formulated if $m = 2$.

It is an immediate consequence of the case $n = m-1$ of (i_n) that every $S \in C^m$ has an isothermic C^m-parametrization; cf. [13], pp. 686-687. In contrast, the "non-euclidean isothermic" parametrization, dealt with by (iii_m), replaces the given C^m-parametrization of S by a C^{m-1}-parametrization, and so one might expect that (iii_m) is not of a final nature. But such is not the case, since the assertion of (iii_m) is as sharp as it can be:

$(\mathrm{iii}_m \text{ bis})$ *If an* S *possesses a* C^{m-1}*-parametrization* (26) *in which the parameter lines* (27) *are asymptotic lines, then* $S \in C^m$.

In fact, this converse, $(\mathrm{iii}_m \text{ bis})$, of (iii_m) is known to follow, for every

$m > 1$, from Codazzi's equations (if $m > 2$, and from the "integrated form" of Codazzi's equations if $m = 2$); cf. [3], p. 139.

8. In order to prove (iii$_m$), consider, for a fixed $m > 1$, an $S \in C^m$ on which the Gaussian curvature does not vanish, and let (1) be the second fundamental form of S with reference to a C^m-parametrization,

$$S: X = X(u^1, u^2), \tag{28}$$

of S. Then it is clear that (1) is a C^{m-2}-metric of non-vanishing $\det g_{ik}$. Let $m > 2$. Then, according to the Appendix of [13], the C^{m-2}-metric (1) is a regular C^{m-2}-metric. In addition, this metric is indefinite if the Gaussian curvature of S is negative. Hence (ii$_{m-2}$) is applicable.

Accordingly, the parametrization of $S \in C^m$ can be chosen in such a way that the second fundamental form becomes a regular C^{m-2}-metric having the normal form (3). But the parameter lines (27) are asymptotic lines of S precisely when the second fundamental form is of the form (3). Hence, in order to conclude (iii$_m$), it is sufficient to appeal to a known general theorem, for which various proofs are available (cf. [13], pp. 681-682, where further references are given, and the finer result contained in (II) of [5], p. 348).*

9. In what follows, the problem of those surface parametrizations (26) will be treated in which the parameter lines (27) are required to be lines of curvature. No restriction will now be placed on the sign of the Gaussian curvature (the latter will even be allowed to change sign on S) but umbilical points will, of course, be excluded from S.

Let (26) be a C^m-parametrization of an $S \in C^m$ and let $m > 1$. Then there exist a first fundamental form

$$E(u,v)\,du^2 + 2F(u,v)\,du\,dv + G(u,v)\,dv^2 \tag{29}$$

and a second fundamental form

$$L(u,v)\,du^2 + 2M(u,v)\,du\,dv + N(u,v)\,dv^2, \tag{30}$$

* On p. 360 of the paper [5] of Dr. Hartman and myself, the passage "Hence, . . . above." is erroneous but can easily be rectified by correcting the first four lines as follows:

"Hence $dx + i\,g\,dy = e^{-i\phi}\omega$ is regular. If (ix) is applied to ω and $R = e^{-i\phi}$ in (u, v)-coordinates, it follows from the choice $g(R) = iR \log R$ that $\phi e^{-i\phi}\omega$ is regular. Since $\phi e^{-i\phi}\omega = \phi(dx + i\,g\,dy)$, the real forms $\phi\,dx$, $\phi g\,dy$ are regular. Consequently, (vii) implies that ϕ_y and $(\phi g)_x$ exist and are continuous. Thus $g > 0$ and the existence and continuity of g_x imply that ϕ is of class C^1."

In line 9, p. 357, of [5], the relation $\gamma = \rho^{-1}$ should be read as follows: $\gamma = \rho^{-1}\epsilon$.

and the coefficients of (29) and (30) are functions of class C^{m-1} and C^{m-2}, respectively (in terms of the given C^m-parametrization (26) of S).

From the coefficients of (29) and (30), the "Dupin form," say

$$A(u,v)du^2 + 2B(u,v)dudv + C(u,v)dv^2, \tag{31}$$

can be obtained, by placing (31) equal to

$$\frac{1}{(EG-F^2)^{\frac{1}{2}}}\begin{vmatrix} dv^2 & -dudv & du^2 \\ E & F & G \\ L & M & N \end{vmatrix} \tag{32}$$

(the factor $1/(EG-F^2)^{\frac{1}{2}}$ is attached to the determinant in (32) in order to make the coefficient matrix of (31) a tensor). At the point (u,v) of S, the Dupin form (32) is identical with the geodesic torsion in the direction (du,dv). The coefficients of (31)-(32) are functions of class C^{m-2} in (u,v), since those of (29) and (30) are.

The definition of the lines of curvature of a surface S shows that the parameter lines (27) of a parametrization (26) of S will be lines of curvature if and only if the coefficients of du^2 and dv^2 in (31)-(32) vanish for all (u,v). (Cf., e. g., [1], pp. 98-99). On the other hand, if (26) is *any* parametrization of S (so that $A \equiv 0$ and $B \equiv 0$ are not assumed), then, since the quadratic form (29) is positive definite, the quadratic form (32) is indefinite at all those points (u,v) at which the proportions

$$L:M:N = E:F:G \tag{33}$$

do not hold. But (33) is equivalent to the definition of an umbilical point (u,v) (provided that a "flat" point (u,v), a point at which all three coefficients of (30) vanish, is considered to be an umbilical point). Hence it is clear that if an S is free of umbilical points, then (ii_n) supplies conditions which assure the existence of certain parametrizations in which the parameter lines are lines of curvature. What is missing in this plan is the result which, in the application of (ii_n) to the problem of asymptotic parametrizations (Section 8), was supplied by the Appendix of [13]. But this will be settled in Sections 11-12.

10. What follows in this manner is the following counterpart (iv_m) of (iii_m):

(iv_m) *If there are no umbilical points on an $S \in C^m$, then S possesses C^{m-1}-parametrizations* (26) *in which the parameter lines* (27) *are lines of curvature of S.*

Since the proof of (iv_m) depends on the case $n = m - 2$ of (ii_n), the situation with regard to the m-values which are allowed in (iv_m) is exactly the same as specified in the comments which follow the wording of (iii_m) in Section 7 above.

The proof of (iv_m) proceeds as follows: If an $S \in C^m$, where $m > 1$, is free of umbilical points, then, in terms of a C^m-parametrization of S, the Dupin form of S is an indefinite C^{m-2}-metric (Section 9). It will be shown in Sections 11-12 that, if $m > 2$, then this C^{m-2}-metric is a regular C^{m-2}-metric. Hence it is clear from Section 9 that (iv_m) follows from the case $n = m - 2$ of (ii_n) in the same way as (iii_m) did.

11. Let an $S \in C^m$, where $m > 2$, be free of umbilical points. Then (32) is a C^{m-2}-metric of non-vanishing determinant (in terms of any C^m-parametrization of S) and has therefore a curvature which, if $m > 3$, is a function of class C^{m-4} (in terms of parameters (u, v) which supply a C^m-parametrization of S). The fact announced in the last paragraph of Section 10 is that this function of class C^{m-4} happens to be of class C^{m-3}, if $m > 2$. This is what will now be verified.

The explicit calculations (which, although they have their geometrical meaning, cannot be avoided) can be simplified by using an isothermic parametrization of S; that is, a parametrization (26) in which (29) reduces to

$$E(u, v)(du^2 + dv^2) \qquad (G = E, F = 0). \tag{34}$$

This assumption involves no loss in the degree of differentiability since, according to [13], p. 687, every $S \in C^m$ has a C^m-parametrization (26) satisfying $G = E$ and $F = 0$. Then the Dupin form (32) simplifies to

$$M\,du^2 + (N - L)\,du\,dv - M\,dv^2. \tag{35}$$

The standard representations of the mean and Gaussian curvature, $H = H(u, v)$ and $K = K(u, v)$, of S reduce to

$$H = \tfrac{1}{2}(L + N)/E, \tag{36}$$

$$K = (LN - M^2)/E^2 \tag{37}$$

(cf., e.g., [1], p. 104, where H is the present $-2H$). Finally, as observed by H. Hopf [6], p. 240, the Codazzi equations can be written in the form of an inhomogeneous Cauchy-Riemann system, as follows:

$$M_v - \tfrac{1}{2}(N - L)_u = EH_u, \qquad \tfrac{1}{2}(L - N)_v - M_u = -EH_v. \tag{38}$$

As in the Appendix of [13], suppose first that $m = 4$. Then the assertion is that, if points (u, v) satisfying (33) are excluded, the (indefinite) C^2-metric

(35) has a curvature which is a function of class C^1 in (u, v). This is what will be verified in Section 12. The verification will be such as to make it clear that the proof applies not only for $m = 4$ but also for $m = 5, 6, \cdots$ and, if Weyl's "integrated form" for the curvature of a regular C^1-metric is applied (cf. [10], pp. 42-44), for $m = 3$ also (cf. the corresponding situation in the Appendix of [13]).

12. First, since (26) is a C^4-parametrization,

$$E(u, v) \in C^3 \qquad (E > 0) \tag{39}$$

holds for the first fundamental form, whereas the coefficients of (36), and therefore those of (35), are functions of class C^2. According to (36)-(37),

$$H(u, v) \in C^2, \tag{40}$$

$$K(u, v) \in C^2. \tag{41}$$

Next, if (31) is any C^2-metric, then the explicit form of the prepresentation of its curvature (cf., e.g., [1], p. 51) shows that the principal part of the curvature, that is, the part in which only the contributions of the second derivatives of the coefficients of (31) are retained, is

$$(A_v - B_u)_v + (C_u - B_v)_u. \tag{41}$$

On the other hand, if (31) is identified with (35) and if only the contribution is retained which, in the preceding sense, corresponds to the "principal part" by virtue of (39), then (41) can be reduced to

$$(M_v - \tfrac{1}{2}(N - L)_u)_v + (-M_u - \tfrac{1}{2}(N - L)_v)_u. \tag{42}$$

Accordingly, the curvature of (35) will be proved to be of class C^1 if it is ascertained that the function (42) is of class C^1. But (38) shows that (42) differs from

$$(H_u)_v - (H_v)_u = H_{uv} - H_{vu} = 0 \tag{43}$$

only in "non-principal" terms (cf. (39)); so that the proof is complete.

13. If not only the "principal part," but all terms, of the curvature of a C^2-metric (31) are taken into account, then the details of the calculation lead, via (36)-(37) and (38), to the result that the curvature of the case (35) of (31) is

$$(H^2 - K)^{-2}\, \partial(H, K)/\partial(u, v) \tag{44}$$

to a trivial factor. The denominator of (44) does not vanish, since (36)-(37) and (34) imply that $H^2 > K$ fails to hold only in the excluded case (33). Clearly, (40)-(41) and (39) imply that (44) is a function of class C^1.

The result (44) can be verified to be equivalent to a classical formula of Weingarten, used by Lie; cf. [1], pp. 234-235. But since Weingarten's proof assumes that the parametrization is in terms of lines of curvature, it leads to a considerable loss in the necessary degree of differentiability. This is the reason why a direct proof had to be given above.

Remark. It is clear that if S has no umbilical points (i.e., if the denominator of (44) does not vanish), then (44) can be a function of class C^{m-3} not only if (26) is a C^m-parametrization (in which case $H \in C^{m-2}$ and $K \in C^{m-2}$). In fact, it is sufficient that (26) be a C^{m-1}-parametrization and that the functions H, K, which are then of class C^{m-3}, be such that their Jacobian

$$\partial(H,K)/\partial(u,v) \tag{45}$$

happens to be a function of class C^{m-2}. Hence, the assumptions of (iv_m) can be generalized so as to involve just the latter conditions.

14. As an application of (iv_m), parametrizations in terms of the particular parameters

$$u^* = X \cdot N, \qquad v^* = X \cdot X, \quad \text{where } X = (x,y,z), \tag{46}$$

will now be dealt with. Here $N = N(u,v)$ is the unit vector of the oriented normal to S; so that $N(u,v) \in C^{m-1}$, if (26) is of class C^m. The parameters (u^*, v^*) on S, defined by (46), are intrinsic from the point of view of euclidean geometry, if the indeterminancy of a translation or a rotation of S in the (x,y,z)-space is disregarded.

The replacement of arbitrary parameters (u,v) by the intrinsic parameters (u^*, v^*) first occurs in a note of Weingarten on minimal surfaces [9] (not listed in [1], p. 541). In what follows, Weingarten's somewhat unsymmetric formulae will be replaced by a more straightforward result, and its deduction will be such as to make clear the nature of the conditions to be imposed on S.

It turns out that it is necessary to make the following two assumptions: S is free of umbilical points (i.e.,

$$H^2 \neq K \tag{47}$$

on S) and S does not have a certain "special" position in the X-space, where $X = (x,y,z)$; cf. (55) below. Note that, whereas (47) is an invariant assumption, the assumption made after (47) refers to the choice of the origin and of the direction of the coordinate axes x, y, z. Such a choice will always

be possible, since S is supposed to be a sufficiently small piece of a surface. (It may be mentioned that the assumptions thus imposed on the position of the surface in the X-space, assumptions expressed by (55) below, correspond to what, in connection with the "median surface" of two surfaces, is a "general position" of the latter; cf. [7], p. 757.)

Finally, (iv_m) will be needed in order to specify the conditions under which the classical equations defining lines of curvature, that is, the equations

$$N_u = -\alpha X_u, \qquad N_v = -\beta X_v \tag{48}$$

(Rodrigues; cf., e.g., [1], pp. 97-98) are valid with a specified degree of differentiability. In (48), the scalars $\alpha = \alpha(u,v)$, $\beta = \beta(u,v)$ denote the principal curvatures of S; so that

$$H = \tfrac{1}{2}(\alpha + \beta), \qquad K = \alpha\beta. \tag{49}$$

15. It will be concluded from (iv_m) that, if an $S \in C^m$ satisfies the conditions mentioned in Section 14, then S possesses a C^{m-1}-parametrization

$$X = X(u^*, v^*) \tag{50}$$

in terms of the "euclidean" parameters (46) of Weingarten. Incidentally, Weingarten used $v^* = |X|$, instead of the $v^* = |X|^2$ of (46); cf. [9] and the rediscovery of (46) in [8] (and, concerning certain results of [8], the comments in [3], pp. 128-129).

First, since $S \in C^m$, it follows from (iv_m) and (47) that S possesses a C^{m-1}-parametrization (26) in which (48) is valid. Next, since the normal N is attached to $S \in C^m$ in a parametrization-invariant manner,

$$N(u,v) \in C^{m-1} \tag{51}$$

(even though only

$$X(u,v) \in C^{m-1} \tag{52}$$

holds; from (52) alone, only $N(u,v) \in C^{m-2}$, rather than (51), would follow).

It is clear from (51), (52) and (46) that both functions

$$u^* = u^*(u,v), \qquad v^* = v^*(u,v) \tag{53}$$

are of class C^{m-1}. Hence, if the Jacobian of (53) does not vanish, then the substitution (53) has an inverse of class C^{m-1}. On the other hand, (50) results if the inverse of (53) is substituted into $X(u,v)$. It follows therefore from (52) that the C^{m-1}-character of (50) will be proved if it is ascertained that the Jacobian of (53) does not vanish.

In view of (46) and (48), the Jacobian matrix of (53) is

$$\begin{pmatrix} u^*_u & u^*_v \\ v^*_u & v^*_v \end{pmatrix} = \begin{pmatrix} -\alpha X \cdot X_u & -\beta X \cdot X_v \\ 2X \cdot X_u & 2X \cdot X_v \end{pmatrix}, \tag{54}$$

since the scalar products $N \cdot X_u$, $N \cdot X_v$ vanish identically (N is normal to (26)). But (49) and (47) imply that $\alpha \neq \beta$. Hence the determinant of (54) does not vanish if

$$X \cdot X_u \neq 0, \qquad X \cdot X_v \neq 0. \tag{55}$$

The pair of conditions (55) is the assumption indicated after (47). Clearly, (55) will be satisfied at every point of any given S if the origin and the direction of the coordinate axes x, y, z in the X-space are chosen in a "non-special" manner (provided that S consists of a sufficiently small neighborhood of any given point of a surface).

16. A parametrization (26) closely related to (50) is

$$S\colon\ X = X(\alpha, \beta), \tag{56}$$

where α, β are the principal curvatures. (56) is substantially equivalent to the parametrization

$$S\colon\ X = X(H; K), \tag{57}$$

provided that umbilical points are excluded from S (i.e., that the sign of equality never holds in the inequality

$$H^2 - K \geqq 0, \tag{58}$$

an inequality which, in view of (49), is always true). In fact, (49) shows that the transition from (56) to (57) is analytic and has therefore an analytic inverse (locally) provided that the Jacobian of (49) does not vanish. But the latter proviso is equivalent to $\alpha \neq \beta$, i.e., to

$$H^2 \neq K, \qquad \text{(hence } H^2 > K\text{)}, \tag{59}$$

since $\alpha = \beta$ is equivalent to

$$H^2 = K. \tag{59_0}$$

On the other hand, the situation with regard to (57) is as follows:

Suppose that $S \in C^m$, where $m > 2$, and let (26) be a C^m-parametrization of S. Then

$$H = H(u, v), \quad K = K(u, v) \text{ are of class } C^{m-2} \qquad (m > 2). \tag{60}$$

Hence, an $S \in C^m$ possesses a C^{m-2}-parametrization (57) whenever the Jacobian (45) does not vanish. In fact, the substitution (60) then possesses an inverse, $(H, K) \rightarrow (u, v)$, which is of class C^{m-2}, and (26) is identical with (57) by virtue of this substitution.

It seems to be peculiar that only

$$\partial(H, K)/\partial(u, v) \neq 0, \tag{61}$$

rather than (61) *and* (59), was used in this deduction. The explanation is that (59_0) cannot hold unless

$$\partial(H, K)/\partial(u, v) = 0. \tag{61_0}$$

17. Let a point (u, v) of an $S \in C^3$ be called a *W-point* if, in some (but then, of course, in every) C^3-parametrization (26) of S, the Jacobian (45) vanishes at that point. By covariance, the same must then hold if (26) is any C^1-parametrization (for instance, any C^2-parametrization) of $S \in C^3$. It is also clear that an $S \in C^3$ is what in a customary sense of the word is a *W-surface* (Weingarten) if and only if every point of S is a W-point.

Every umbilical point of an $S \in C^3$ is a W-point. In fact, a comparison of (58) with (59_0) shows that, if (u, v) is an umbilical point, then the function $H^2 - K$ acquires a (local, not necessarily isolated) minimum and has, therefore, a stationary value. But then the gradient of $H^2 - K$ vanishes. This means that the vector $(2H, -1)$ is a solution of two homogeneous, linear equations having the determinant (45), which proves that (61_0) is satisfied.

The converse of this deduction is not possible, i. e., a W-point need not be a stationary point of the function $H^2 - K$. In fact, the latter is known to be stationary at every point of (that is, constant on) an $S \in C^3$ only if S is a plane, a sphere or a circular cylinder, whereas (45) will vanish identically on every surface of revolution.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] L. Bianchi, *Lezioni di geometria differenziale*, 1st ed. (1894).

[2] S.-S. Chern, P. Hartman and A. Wintner, "On isothermic coordinates," *Commentarii Mathematici Helvetici*, vol. 28 (1954), pp. 301-309.

[3] P. Hartman and A. Wintner, "On the curvatures of a surface," *American Journal of Mathematics*, vol. 75 (1953), pp. 127-141.

[4] ——— and A. Wintner, "On asymptotic parametrizations," *ibid.*, vol. 75 (1953), pp. 486-496.

[5] ——— and A. Wintner, "Regular binary Pfaffians and non-parabolic partial differential equations," *Rendiconti del Circolo Matematico di Palermo*, ser. 2, vol. 3 (1955), pp. 347-362.

[6] H. Hopf, "Über Flächen mit einer Relation zwischen den Hauptkrümmungen," *Mathematische Nachrichten*, vol. 4 (1951), pp. 232-249.

[7] ——— and H. Samelson, "Zum Beweis des Kongruensatzes für Eiflächen," *Mathematische Zeitschrift*, vol. 43 (1938), pp. 749-766.

[8] W. Scherrer, "Stützfunktion und Radius. I.," *Commentarii Matematici Helvetici*, vol. 20 (1947), pp. 366-381.

[9] J. Weingarten, "Eine neue Klasse auf einander abwickelbarer Flächen," *Göttinger Nachrichten* (1887), pp. 28-31.

[10] H. Weyl, "Ueber die Bestimmung einer geschlossenen konvexen Fläche durch ihr Linienelement," Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich, vol. 61 (1915), pp. 40-72.

[11] A. Wintner, "On isometric surfaces," *American Journal of Mathematics*, vol. 74 (1952), pp. 198-214.

[12] ———, "On the infinitesimal geometry of curves," *ibid.*, vol. 75 (1953), pp. 241-259.

[13] ———, "On the local rôle of the theory of the logarithmic potential in differential geometry," *ibid.*, vol. 75 (1953), pp. 679-690.

[14] ———, "On the local embedding problems in the differential geometry of surfaces," vol. 77 (1955), pp. 845-852.

ON THE GALOIS THEORY OF DIFFERENTIAL FIELDS.*

By E. R. KOLCHIN.[1]

Introduction.

1. Summary. In a preceding paper [7] there was presented a Galois theory, for a certain kind of differential field extension called strongly normal. The Galois group of a strongly normal extension is endowed with a structure very much like that of a group variety, as studied by Weil [14]. In the present paper the study of such Galois groups is renewed for the purpose of clarifying their connection with group varieties on the one hand and with strongly normal extensions on the other.

Consider a differential field $\mathfrak{F}$ of characteristic 0 with algebraically closed field of constants $\mathfrak{C}$, and a strongly normal extension $\mathfrak{G}$ of $\mathfrak{F}$; suppose given a universal extension $\mathfrak{F}^*$ of $\mathfrak{G}$, and denote the field of constants of $\mathfrak{F}^*$ by $\mathfrak{C}^*$. For reasons of convenience we use the term "Galois group of $\mathfrak{G}$ over $\mathfrak{F}$" for the group of all (ipso facto strong) isomorphisms of $\mathfrak{G}$ over $\mathfrak{F}$, and not for its subgroup consisting of all automorphisms of $\mathfrak{G}$ over $\mathfrak{F}$. The field $\mathfrak{C}^*$, being algebraically closed and of finite transcendence degree over $\mathfrak{C}$, may be used as a universal domain for algebraic geometry (Weil [13], ch. I, § 1); it is the group varieties of this algebraic geometry, defined over $\mathfrak{C}$ and slightly generalized to permit group varieties which are reducible (i.e. which have more than one component), which we consider. By an "algebraic group" we mean either a Galois group or a group variety as above.

Chapter I is primarily a study of "rational" homomorphisms of algebraic groups into algebraic groups; these seem to be the homomorphisms appropriate to the consideration of algebraic groups. Rational homomorphisms of group varities into group varieties, without restriction to fields of characteristic 0, were considered by Weil [14] (who omitted the adjective "rational").

Specializing the concept of rational homomorphism we obtain that of birational isomorphism. In Chapter II it is shown that every Galois group is birationally isomorphic to a group variety, and conversely that every *irreducible* group variety is birationally isomorphic to a Galois group; this

* Received June 29, 1955.

[1] Some of the results in this paper were obtained in connection with a contract with the National Science Foundation, and some while the author was a Guggenheim fellow.

answers a question raised in [7], p. 759. The first of these results has already been proved by Matsumura [8] for irreducible Galois groups (the extension to reducible Galois groups is fairly trivial); the present proof, although shorter, makes use of a special theorem due to Weil and refined by Rosenlicht (see Chapter II, § 1 below), whereas Matsumura's proof has the merit of being more self-contained. The proof of the second of these results makes use of the fact that the vector space of derivations of the field of rational functions on an irreducible group variety has a base consisting of invariant derivations (see Chapter II, § 2 below); this fact, which is central in the proof, was pointed out to me in this connection by Chevalley. Whether or not the result can be extended to reducible group varieties remains open. Chapter II also contains certain consequences of these two results.

In Chapter III the concept of a rational homomorphism of the Galois group $\mathfrak{G}$ of $\mathfrak{H}$ over $\mathfrak{F}$ into a group variety G is generalized to obtain that of a rational crossed homomorphism, or a cocycle of dimension 1, or a 1-cocycle, of $\mathfrak{G}$ into G. This leads, in much the usual way, to the set $H^1(\mathfrak{G}, G)$ of cohomology classes of dimension 1 of $\mathfrak{G}$ into G. Since G is not assumed to be commutative, $H^1(\mathfrak{G}, G)$ is not in general a group. It is proved that $H^1(\mathfrak{G}, G)$ is trivial provided *either* G is a full matric group *or* $\mathfrak{F}$ is algebraically closed; also, if G is commutative (so that $H^1(\mathfrak{G}, G)$ is a group) then every element of $H^1(\mathfrak{G}, G)$ has finite order. Using the first of these results, it is shown that the strongly normal extension $\mathfrak{H}$ of $\mathfrak{F}$ is a Picard-Vessiot extension if and only if its Galois group $\mathfrak{G}$ is birationally isomorphic to an algebraic group of matrices; this answers another question raised in [7], p. 759. Similarly, the last of these results is used to give a short proof of the characterization given in [7] of all strongly normal extensions of transcendence degree 1.

2. Assumptions and notation. Throughout this paper $\mathfrak{F}$ denotes a differential field of characteristic 0 with derivation operators $\delta_1, \cdots, \delta_m$ and algebraically closed field of constants $\mathfrak{k}$, and $\mathfrak{H}$ denotes a strongly normal extension of $\mathfrak{F}$. We fix a universal extension $\mathfrak{F}^*$ of $\mathfrak{H}$ and denote the field of constants of $\mathfrak{F}^*$ by $\mathfrak{k}^*$; except when the contrary is explicitly stated, all differential fields considered are differential subfields of $\mathfrak{F}^*$, so that every isomorphism of a differential field has image in $\mathfrak{F}^*$.

Because $\mathfrak{H}$ is strongly normal over $\mathfrak{F}$, every isomorphism σ of $\mathfrak{H}$ over $\mathfrak{F}$ is strong; denoting the field of constants of $\mathfrak{H}\langle\sigma\mathfrak{H}\rangle$ by $\mathfrak{k}(\sigma)$ (instead of by $\mathfrak{k}_\sigma$ as in [7]), we can express this fact by writing

$$\mathfrak{H}\langle\sigma\mathfrak{H}\rangle = \mathfrak{H}\langle\mathfrak{k}(\sigma)\rangle = (\sigma\mathfrak{H})\langle\mathfrak{k}(\sigma)\rangle.$$

As in [7], we identify each such σ with the unique automorphism of $\mathfrak{H}\langle\mathfrak{k}^*\rangle$ over $\mathfrak{F}\langle\mathfrak{k}^*\rangle$ which extends σ; the set of all isomorphisms of $\mathfrak{H}$ over $\mathfrak{F}$ is thereby identified with the group of all automorphisms of $\mathfrak{H}\langle\mathfrak{k}^*\rangle$ over $\mathfrak{F}\langle\mathfrak{k}^*\rangle$. We call this group the *Galois group* of $\mathfrak{H}$ over $\mathfrak{F}$ and denote it by $\mathfrak{G}$ (instead of by $\mathfrak{G}^*$ as in [7]); we recall that $\mathfrak{G}$ is algebraic in the sense defined in [7]. The group of all automorphisms of $\mathfrak{H}$ over $\mathfrak{F}$ is a subgroup of $\mathfrak{G}$. If $\sigma_1, \cdots, \sigma_p \in \mathfrak{G}$ we denote by $\mathfrak{k}(\sigma_1, \cdots, \sigma_p)$ the field of constants of $\mathfrak{H}\langle\sigma_1\mathfrak{H}, \cdots, \sigma_p\mathfrak{H}\rangle$, that is, the compositum of the p fields $\mathfrak{k}(\sigma_1), \cdots, \mathfrak{k}(\sigma_p)$. It is easy to see that $\mathfrak{k}(\sigma^{-1}) = \mathfrak{k}(\sigma)$ and $\mathfrak{k}(\sigma\tau, \sigma) = \mathfrak{k}(\sigma, \tau) = \mathfrak{k}(\sigma\tau, \tau)$ for all $\sigma, \tau \in \mathfrak{G}$.

Chapter I. Rational homomorphisms.

1. Varieties. The field $\mathfrak{k}^*$ is an algebraically closed extension of $\mathfrak{k}$ of infinite transcendence degree, and therefore may be used as a universal domain for algebraic geometry. We find it convenient to deviate from Weil's terminology, in order to permit varieties which are not irreducible. Specifically, we adopt the following terminology, which is sufficient for our purpose:[2] we call *irreducible variety* what Weil calls abstract variety; we call *variety* any finite union of irreducible subvarieties of some irreducible variety (in Weil's terminology, a bunch of subvarieties of an abstract variety), or the union of a disjoint finite set of such. Every irreducible variety of dimension > 0 which we consider, except those for which the contrary is explicitly stated, is assumed to have $\mathfrak{k}$ as a field of definition. Similarly, when we speak of a specialization of a point it is always to be understood, except when the contrary is stated, that the specialization is over $\mathfrak{k}$. We find it convenient sometimes to call a subset of a variety *algebraic* to indicate that the subset is a subvariety. What we mean by the *components* of a variety is clear.

Let V and W be varieties, and denote the components of V by $V_1, \cdots, V_h$.

[2] An intrinsic definition of abstract algebraic variety, encompassing the case of reducible varieties, has been given by Serre [12]. The present definition, which involves an imbedding in some irreducible variety, is adopted here in the interest of brevity. This is, of course, merely a difference of procedure and not of substance. However, in the discussion below of rational mappings, the definition adopted for what it means for a rational mapping to be defined at a particular point is weaker than the definition that naturally arises in developing farther the ideas of Serre. Specifically, the present definition leads to a notion of isomorphism of varieties such that two components of a variety can have different higher intersection properties from the corresponding components of an isomorphic variety; on the other hand, because tthe definition here is weaker, Proposition 1 below is stronger. Of course, when dealing with varieties for which the components are disjoint (which by Proposition 1 is the case for group varieties) this difference in the two definitions disappears.

By a *rational mapping* of V into W we mean any sequence $f = (f_1, \cdots, f_h)$ where, for each i, f_i is a rational mapping of V_i into some component of W (that is, f_i is what is called by Weil [14] a function defined on V_i with values in a component of W); by a *field of definition* of f we mean a common field of definition of $f_1, \cdots, f_h$. All rational mappings except those for which the contrary is stated, are assumed to have $\mathfrak{k}$ as a field of definition. We say the rational mapping f is *defined* at a point $s \in V$ if, for every point $t \in V$ of which s is a specialization, there exists an index i such that $t \in V_i$ and f_i is defined at t, and for all such indices i the value $f_i(t)$ is the same. If f is defined at s then there exists an i such that f_i is defined at s, and the point $f_i(s)$ is independent of the choice of i; we call this point the *value* of f at s and denote it by $f(s)$. This definition insures that if s' is a specialization of $s \in V$ and f is defined at s', then f is defined at s and $(s', f(s'))$ is a specialization of $(s, f(s))$. If ϕ is a mapping (in the set-theoretic sense) of a subset of V into W, then we permit ourselves to say that ϕ is *rational* if there exists a rational mapping f of V into W such that f is defined at precisely those points $s \in V$ at which ϕ is defined and $f(s) = \phi(s)$ at all such points s; in other words, we identify each rational mapping of V into W with a mapping, which we call rational.

By a *biregular birational mapping of V onto W* we mean a one-to-one mapping f of V onto W such that f and f^{-1} are both rational and everywhere defined.

By a *group variety* we mean any variety G with a group structure such that (when the group composition is written multiplicatively) the mapping $(s, s') \to ss'$ of $G \times G$ into G and the mapping $s \to s^{-1}$ of G into G are rational. Using an argument which is familiar in several special cases and analogous situations ([5], § 2, Th., p. 10; Weil [14], § 4, Prop. 9, pp. 39-40; Chevalley [3], Ch. I, § 3, Th. 2, p. 86; [7], Ch. II, Th. 1, p. 788), it is easy to prove the following result.

PROPOSITION 1. *Precisely one component of any group variety G contains the unity element; this component G^0 is a normal subgroup of finite index, and the cosets of G relative to G^0 are the components of G.*

Thus, the components of G are disjoint and all have the same dimension, which we call the dimension of G and denote by $\dim G$.

PROPOSITION 2. *If $\mathfrak{h}$ is a subgroup of the group variety G consisting solely of points rational over $\mathfrak{k}$, and H is the smallest subvariety of G containing $\mathfrak{h}$, then H is a subgroup of G.*

Proof. For any $s \in \mathfrak{h}$, sH is the smallest subvariety of G containing $s\mathfrak{h} = \mathfrak{h}$, so that $sH = H$; therefore $\mathfrak{h}H = H$. For every $t \in H$ which is rational over $\mathfrak{k}$, Ht is the smallest subvariety of G containing $\mathfrak{h}t$, and $\mathfrak{h}t \subset \mathfrak{h}H = H$, so that $Ht \subset H$; this inclusion obviously continues to hold without the restriction that t be rational over $\mathfrak{k}$, so that $HH \subset H$. Finally, H^{-1} is the smallest subvariety of G containing $\mathfrak{h}^{-1} = \mathfrak{h}$, so that $H^{-1} = H$. Therefore H is a group.

PROPOSITION 3. *If $\mathfrak{g}$ is a subgroup of the group variety G such that $G - \mathfrak{g}$ is contained in a subvariety of G all the components of which are of lower dimension than G, then $\mathfrak{g} = G$.*

Proof. If $\mathfrak{g}$ were different from G, a coset $s\mathfrak{g}$ of G relative to $\mathfrak{g}$ would be contained in $G - \mathfrak{g}$ and therefore in a subvariety of G all the components of which have dimension $< \dim G$, and we obviously may suppose s rational over $\mathfrak{k}$; therefore $\mathfrak{g}$ would be contained in such a variety, so that $G = \mathfrak{g} \cup (G - \mathfrak{g})$ would, too, contradiction.

The proofs of Propositions 2 and 3 are the same as those of the analogous propositions on algebraic groups of strong isomorphisms of a differential field (see [7], Ch. II, Props. 14, 15, p. 789, for the case of algebraic groups of automorphisms).

2. Rational homomorphisms of group varieties. Consider two group varieties G and H, and a rational mapping f of G into H such that $f(st) = f(s)f(t)$ whenever s and t are points of G such that f is defined at s, t, and st (or what amounts to the same thing, whenever s and t are independent generic points of components of G). We call such a rational mapping a *rational homomorphism* of G into H.

PROPOSITION 4. *A necessary and sufficient condition that f be a rational homomorphism of the group variety G into the group variety H is that f be a homomorphism of the group G into the group H having the two properties:* 1) *for all $s \in G$, $\mathfrak{k}(f(s)) \subset \mathfrak{k}(s)$;* 2) *for all s, $s' \in G$ with s' a specialization of s, $f(s')$ is a specialization of $f(s)$.*

Proof. Let f be a rational homomorphism of G into H, and let s' be any point of G. Choose independent generic points s, t of components of G such that s' is a specialization of s; there exists a specialization t' of t such that t' is rational over $\mathfrak{k}$ and f is defined at t' and $s't'$. Then f is defined at st', and $f(s) = f(st')f(t')^{-1}$. This makes it clear that f is defined at s', that is, f is defined at every point of G. Thus, f is a homomorphism (in the

group theoretic sense) of the group G into the group H, and f obviously has the two required properties.

Conversely, let f be any homomorphism of the group G into the group H having these two properties. Letting $s_1, \cdots, s_n$ be independent generic points of the n components of G, we see from 1) that there exists a rational mapping g of G into H such that $g(s_i) = f(s_i)$ $(1 \leqq i \leqq h)$; we then see from 2) that $g(s) = f(s)$ for every generic point of every component of G, so that $g(st) = g(s)g(t)$ whenever s and t are independent generic points of components of G; in other words, g is a rational homomorphism of G into H. Therefore g is everywhere defined and, by 2), $g(s) = f(s)$ for every $s \in G$.

PROPOSITION 5. *Let f be a rational homomorphism of a group variety G into a group variety H. If N is any algebraic set in H then $f^{-1}(N)$ is an algebraic set in G. If M is an algebraic subgroup of G then $f(M)$ is an algebraic subgroup of H.*

Proof. Let A be the smallest algebraic set in G containing $f^{-1}(N)$, let A_0 be any component of A, and let s be a generic point of A_0. If $f(s)$ were not a point of N there would exist a proper algebraic subset A_1 of A_0 such that every specialization s' of s with $s' \notin A_1$ has the property that $f(s') \notin N$, that is $f^{-1}(N)$ would be contained in the algebraic set obtained from A on replacing the component A_0 by A_1, so that A would not be the smallest alegbraic set in G containing $f^{-1}(N)$. Therefore $f(s) \in N$, so that $f(A_0) \subset N$, whence $f(A) \subset N$. Thus $f^{-1}(N) = A$, which is algebraic.

Let B be the smallest algebraic set in H containing $f(M)$; obviously B is also the smallest algebraic set in H containing $f(M')$, where M' is the group of all points of M which are rational over $\mathfrak{k}$, so that by Proposition 2 B is an algebraic group. Let $s_1, \cdots, s_h$ be generic points of all the components of M, and let B_i be the irreducible algebraic set in H with generic point $f(s_i)$. Clearly $B = B_1 \cup \cdots \cup B_h$. For each i there exists a proper algebraic subset C_i of B_i such that every specialization t_i' of $f(s_i)$ with $t_i' \notin C_i$ has the property that $t_i' = f(s_i')$ for some specialization s_i' of s_i. Thus, $B - f(M)$ is contained in an algebraic set every component of which is a proper subset of a component of B. It follows from Proposition 3 that $f(M) = B$.

COROLLARY. *The kernel and the image of a rational homomorphism are algebraic.*

We call a rational homomorphism of a group variety G into a group

variety H a *birational isomorphism* of G onto H provided the kernel is trivial, the image is H, and the inverse is a rational homomorphism of H into G. It follows from Proposition 4 that f is a birational isomorphism of G onto H if and only if f is an isomorphism of the group G onto the group H such that: 1) $\mathscr{C}(s) = \mathscr{C}(f(s))$ for every $s \in G$; 2) for all s, s' in G, s' is a specialization of s if and only if $f(s')$ is a specialization of $f(s)$.

3. Rational homomorphisms in general. Henceforth we use the term "algebraic group" to mean *either* a Galois group (of a strongly normal extension of a differential field with field of constants $\mathscr{C}$), *or else* a group variety. We generalize to arbitrary algebraic groups the notion of rational homomorphism previously defined for group varieties.

By a *rational homomorphism* of an algebraic group G into an algebraic group H we mean any homomorphism f of the group G into the group H having the two properties that whenever $s \in G$ then $\mathscr{C}(f(s)) \subset \mathscr{C}(s)$, and whenever s, $s' \in G$ and s' is a specialization of s then $f(s')$ is a specialization of $f(s)$. If in addition the kernel of f is trivial and the image of f is H, so that f is an isomorphism of the group G onto the group H, and if f^{-1} is a rational homomorphism of H into G, then we say that f is a *birational isomorphism* of G onto H. In other words, a birational isomorphism of G onto H is an isomorphism f of the group G onto the group H such that $\mathscr{C}(f(s)) = \mathscr{C}(s)$ $(s \in G)$ and s' is a specialization of s if and only if $f(s')$ is a specialization $f(s)$ $(s, s' \in G)$.

Example 1. Let $\mathscr{U}$ be a differential field between $\mathscr{F}$ and $\mathscr{G}$ which is strongly normal over $\mathscr{F}$, and denote the Galois group of $\mathscr{U}$ over $\mathscr{F}$ by $\mathfrak{H}$. We know ([7], Ch. III, Th. 4, p. 797) that the algebraic subgroup $\mathfrak{G}(\mathscr{U})$ of $\mathfrak{G}$ corresponding to the intermediate differential field $\mathscr{U}$ is normal, and that the mapping r which to each $\sigma \in \mathfrak{G}$ corresponds the restriction of σ to $\mathscr{U}$ is a homomorphism, with kernel $\mathfrak{G}(\mathscr{U})$ and image $\mathfrak{H}$, of the group $\mathfrak{G}$ into the group $\mathfrak{H}$. For every $\sigma \in \mathfrak{G}$, the field of constants of $\mathscr{U}\langle r(\sigma)\mathscr{U}\rangle = \mathscr{U}\langle \sigma\mathscr{U}\rangle$ is obviously contained in the field of constants of $\mathscr{G}\langle \sigma\mathscr{G}\rangle$, that is, $\mathscr{C}(r(\sigma)) \subset \mathscr{C}(\sigma)$. Also, if $(\sigma'\alpha)_{\alpha \in \mathscr{G}}$ is a specialization of $(\sigma\alpha)_{\alpha \in \mathscr{G}}$ over $\mathscr{G}$ then $(r(\sigma')\alpha)_{\alpha \in \mathscr{U}} = (\sigma'\alpha)_{\alpha \in \mathscr{U}}$ is a specialization of $(r(\sigma)\alpha)_{\alpha \in \mathscr{U}} = (\sigma\alpha)_{\alpha \in \mathscr{U}}$ over $\mathscr{U}$, that is, if σ' is a specialization of σ then $r(\sigma')$ is a specialization of $r(\sigma)$. Therefore r is a rational homomorphism of $\mathfrak{G}$ into $\mathfrak{H}$.

Example 2. If $\mathscr{G}$ is a Picard-Vessiot extension of $\mathscr{F}$ then ([5] and [6]) $\mathscr{G}$ has a fundamental system of generators, that is, a finite number of elements

$\eta_1, \cdots, \eta_n$ linearly independent over constants such that $\mathfrak{H} = \mathfrak{F}\langle \eta_1, \cdots, \eta_n \rangle$ and such that for every $\sigma \in \mathfrak{G}$ there exists an invertible matrix of constants

$$c(\sigma) = (c_{ij}(\sigma))_{1 \leqq i \leqq n,\ 1 \leqq j \leqq n}$$

for which

$$\sigma \eta_j = \sum_{i=1}^{n} c_{ij}(\sigma) \eta_i \qquad (1 \leqq j \leqq n).$$

The mapping $c : \sigma \to c(\sigma)$ $(\sigma \in \mathfrak{G})$ is an isomorphism of $\mathfrak{G}$ onto an algebraic subgroup of the group variety $GL(n)$ of all invertible square matrices of degree n with coordinates in $\mathfrak{k}^*$. It is not difficult to see that c is a birational isomorphism.

It is obvious that if f is a rational homomorphism of an algebraic group G into an algebraic group H and g is a rational homomorphism of H into an algebraic group I, then $g \circ f$ is a rational homomorphism of G into I.

PROPOSITION 6. *Let f and f' be rational homomorphisms of an algebraic group G into group varieties H and H' respectively. If f and f' have the same kernel and have respective images H and H', then the isomorphism ϕ of H onto H' such that $\phi \circ f = f'$ is birational.*

Proof. Letting s be any element of G we see that for a generic specialization s' of s, $f(s) = f(s')$ if and only if $f'(s) = f'(s')$; this means that an isomorphism of $\mathfrak{k}(s)$ leaves invariant each element of $\mathfrak{k}(f(s))$ if and only if it leaves invariant each element of $\mathfrak{k}(f'(s))$, so that $\mathfrak{k}(f(s)) = \mathfrak{k}(f'(s))$; thus $\mathfrak{k}(t) = \mathfrak{k}(\phi(t))$ for every $t \in H$. Now, if s is a generic point of a component of G then $f(s)$ is a generic point of a component of H, and every generic point of a component of H is of this form; similarly for H'. If t' is a generic specialization of $t = f(s)$, then there exists a generic specialization s' of s such that $t' = f(s')$, and obviously $f'(s')$ is a generic specialization of $f'(s)$. Therefore there exists a rational mapping g of H into H' such that $g(t) = \phi(t)$ for every generic point t of a component of H. If t, t' are independent generic points of components of H then we may write $t = f(s)$, $t' = f(s')$, where s, s' are independent generic elements of components of G, so that $g(tt') = g(f(s)f(s')) = g(f(ss')) = \phi(f(ss')) = f'(ss') = f'(s)f'(s') = \phi(t)\phi(t') = g(t)g(t')$. Therefore g is a rational homomorphism of H into H'. It follows that $g \circ f$ is a rational homomorphism of G into H' coinciding with f' at every generic point of every component of G and therefore at every point whatever, that is $g \circ f = f'$. Therefore $\phi = g$, so that ϕ is a rational homomorphism of H into H'.

Reversing the roles of f and f' we see that ϕ^{-1} is rational too, so that ϕ is birational.

COROLLARY. *If f is a rational homomorphism of a group variety G into a group variety H such that the kernel of f is trivial and the image of f is H, then f is a birational isomorphism of G onto H.*

Proof. In the proposition set $H' = G$, $f' =$ the identity mapping of G.

PROPOSITION 7. *Let G be an algebraic group, G_1 a normal algebraic subgroup of G of finite index, and K any normal algebraic subgroup of G. Then every rational homomorphism f_1 of G_1 with kernel $G_1 \cap K$ and image a group variety can be extended to a rational homomorphism f of G with kernel K and image a group variety.*

Proof. We extend f_1 in two stages; first to a rational homomorphism of G_1K with kernel K and the same image as f_1, and then to a rational homomorphism of G. G_1K is the union of its cosets relative to G_1; these are finite in number and each is an algebraic set, so that G_1K is a normal algebraic subgroup of G. Every element s of G_1K is of the form $s = s_0a$, where $s_0 \in G_1$ and $a \in K$, and $f_1(s_0)$ clearly depends only on s and not on the particular choice of s_0 and a. Therefore we may define a mapping g of G_1K into the image H_1 of f_1 by the formula $g(s_0a) = f_1(s_0)$. The computation $g(s_0'a's_0a) = g(s_0's_0 \cdot s_0^{-1}a's_0a) = f_1(s_0's_0) = f_1(s_0')f_1(s_0) = g(s_0'a')g(s_0a)$ shows that g is a homomorphism of G_1K with image H_1, and clearly g has kernel K. Now every coset of G_1K relative to G_1 contains an element of K, and therefore an element $a \in K$ such that $\mathfrak{B}(a) = \mathfrak{B}$. Therefore every element $s \in G_1K$ can be written in the form $s = s_0a$ with $s_0 \in G_1$, $a \in K$, and $\mathfrak{B}(a) = \mathfrak{B}$, so that $\mathfrak{B}(g(s)) = \mathfrak{B}(f(s_0)) \subset \mathfrak{B}(s_0) = \mathfrak{B}(s_0a) = \mathfrak{B}(s)$; furthermore, if s' is a specialization of s then s' belongs to the same coset as s, so that $s' = s_0'a$, where $s_0' \in G_1$; obviously s_0' is a specialization of s_0, so that $f_1(s_0')$ is a specialization of $f_1(s_0)$, whence $g(s')$ is a specialization of $g(s)$. Thus, the homomorphism g is rational.

Let $s_1, \cdots, s_r$ be a complete set of representatives of the cosets of G relative to G_1K; we assume without loss of generality that $\mathfrak{B}(s_i) = \mathfrak{B}$ $(1 \leqq i \leqq r)$ and s_1 is the unity element of G. For each index i the mapping ϕ_i of G_1K defined by $\phi_i(s) = s_i s s_i^{-1}$ is clearly a birational isomorphism of G_1K onto G_1K. For each pair of indices i, j there is an index $k(i,j)$ and an element $s(i,j) \in G_1K$ such that $s_i s_j = s(i,j)s_{k(i,j)}$. Let $H_2, \cdots, H_r$ be varieties biregularly birationally equivalent to H_1 such that $H_1, H_2, \cdots, H_r$ are disjoint, and let h_i be a biregular birational mapping of H_1 into H_i $(1 \leqq i \leqq r)$, where h_1 is the identity mapping of H_1.

Since g and $g \circ \phi_i$ are rational homomorphisms of G_1K with the same kernel K and image H_1, there exists by Proposition 6 a birational isomorphism ψ_i of H_1 onto H_1 such that $\psi_i \circ g = g \circ \phi_i$. Furthermore, the mapping ρ_{ij} of H_1 defined by $\rho_{ij}(t) = t \cdot g(s(i,j))$ is a biregular birational mapping of H_1 onto H_1. Also, the law of composition, which we denote by μ_1, of the group variety H_1 is an everywhere defined rational mapping of $H_1 \times H_1$ into H_1. Consequently the mapping

$$\mu_{ij} = h_{k(i,j)} \circ \rho_{ij} \circ \mu_1 \circ (h_i^{-1} \times (\psi_i \circ h_j^{-1}))$$

is an everywhere defined rational mapping of $H_i \times H_j$ into the variety $H = H_1 \cup \cdots \cup H_r$, and it is easy to see that $\mu_{11} = \mu_1$. There is a unique everywhere defined rational law of composition μ on H which extends every μ_{ij}; since μ extends μ_1 we may write $\mu(t, t') = tt'$ for $t, t' \in H$.

Let f be the mapping of G into H defined by $f(ss_i) = h_i(g(s))$ $(s \in G_1K, 1 \leqq i \leqq r)$. Clearly f has image H, extends f_1, and because

$$\begin{aligned} f(ss_i \cdot s's_j) &= f(s \cdot s_i s' s_i^{-1} \cdot s_i s_j) = f(s\phi_i(s')s(i,j)s_{k(i,j)}) \\ &= h_{k(i,j)}(g(s)g(\phi_i(s'))g(s(i,j))) \\ &= (h_{k(i,j)} \circ \rho_{ij} \circ \mu_1 \circ (h_i^{-1} \times (\psi_i \circ h_j)))(h_i(g(s)), h_j(g(s'))) \\ &= h_i(g(s))h_j(g(s')) = f(ss_i)f(s's_j), \end{aligned}$$

f is a homomorphism of the group G. Therefore H, with the law of composition μ, is a group. To prove that H is a group variety it remains to show that $t \to t^{-1}$ $(t \in H)$ is an everywhere defined rational mapping of H into H. Now, the mappings λ_i and ρ_i of H defined by $\lambda_i(t) = f(s_i)^{-1}t$ and $\rho_i(t) = tf(s_i)^{-1}$ are everywhere defined rational mappings of H into H mapping H_i into H_1, and the mapping ν of H_1 defined by $\nu(t) = t^{-1}$ is an everywhere defined rational mapping of H_1 into H; but the mapping $t \to t$ $(t \in H)$ obviously coincides with $\lambda_i \circ \nu \circ \rho_i$ on H_i and therefore is rational and everywhere defined.

Finally, if $s \in G_1K$ then

$$\mathfrak{G}(f(ss_i)) = \mathfrak{G}(h_i(g(s))) = \mathfrak{G}(g(s)) \subset \mathfrak{G}(s) = \mathfrak{G}(ss_i),$$

and if $s's_i$ is a specialization of ss_i then s' is a specialization of s, $g(s')$ is a specialization of $g(s)$, and $f(s's_i) = h_i(g(s'))$ is a specialization of $f(ss_i) = h_i(g(s))$, so that f is rational.

COROLLARY. *A birational isomorphism of the component of the identity $\mathfrak{G}^0$ of the Galois group $\mathfrak{G}$ onto a group variety can always be extended to a birational isomorphism of $\mathfrak{G}$ onto a group variety.*

Proof. In the proposition set $G = \mathfrak{G}$, $G_1 = \mathfrak{G}^0$, and $K =$ the group consisting of the identity automorphism.

Chapter II. Birational isomorphisms between Galois groups and group varieties.

1. Galois groups onto group varieties. In this section we shall prove the following theorem.

THEOREM 1. *The Galois group $\mathfrak{G}$ of the strongly normal extension $\mathfrak{H}$ of the differential field $\mathfrak{F}$ of characteristic 0 with algebraically closed field of constants $\mathfrak{k}$ is always birationally isomorphic to a group variety defined over $\mathfrak{k}$.*

Proof. By the corollary to Proposition 7 of Chapter I, we may assume that $\mathfrak{G}$ is irreducible. To prove the theorem under this added hypothesis, we use a theorem due to Weil ([14], Th. 15, p. 54), and later refined by Rosenlicht ([11], Th. 4, p. 511). To state this theorem we need the following definition due to Weil ([14], p. 51): If W is any irreducible variety, a *normal law of composition* on W is a rational mapping of $W \times W$ into W which, if we denote the law of composition multiplicatively, has the two properties that 1) $\mathfrak{k}(s, st) = \mathfrak{k}(t, st) = \mathfrak{k}(t, s)$ for independent generic points s and t of W, and 2) $(st)u\ s(tu)$ for independent generic points s, t, and u of W. Weil's theorem is: If W is an irreducible variety with a normal law of composition, then there exists a birational correspondence T between W and a group variety G which transforms the normal law of composition on W into the group composition on G (T, G, and the group composition on G not necessarily defined over $\mathfrak{k}$). Rosenlicht's contribution is to add: If the points of W which are rational over $\mathfrak{k}$ form a dense (Zariski topology) subset of W (which is true in the present case because $\mathfrak{k}$ is algebraically closed), then we may take G and T so that T, G, and the group composition on G are defined over $\mathfrak{k}$.

Let σ, τ be generic elements of $\mathfrak{G}$ which are independent ([7], pp. 777-778); then $\sigma\tau$ is also a generic element of $\mathfrak{G}$. Let $\gamma_1, \cdots, \gamma_r$ be constants such that $\mathfrak{k}(\sigma) = \mathfrak{k}(\gamma_1, \cdots, \gamma_r)$. Each γ_i can be expressed as the quotient of two polynomials over $\mathfrak{H}$ in a finite number of elements $\sigma\eta_1, \cdots, \sigma\eta_n$, where each $\eta_j \in \mathfrak{H}$; fixing two such polynomials, for each generic element ρ of $\mathfrak{G}$ let $c_i(\rho)$ be the element obtained on replacing each $\sigma\eta_j$ by $\rho\eta_j$ in the quotient of the two polynomials. Then $c_i(\sigma) = \gamma_i$, and for each generic element ρ of $\mathfrak{G}$, $\mathfrak{k}(\rho) = \mathfrak{k}(c_1(\rho), \cdots, c_r(\rho))$. Also, writing $c(\rho) = (c_1(\rho), \cdots, c_r(\rho))$, we see that $c(\sigma)$ is a generic point of a certain variety W in r-dimensional affine space, that $(c(\sigma), c(\tau))$ is a generic point of $W \times W$, and that $c(\sigma\tau)$ is a generic point of W with $\mathfrak{k}(c(\sigma\tau)) = \mathfrak{k}(\sigma\tau) \subset \mathfrak{k}(\sigma, \tau) = \mathfrak{k}(c(\sigma), c(\tau))$.

Therefore there exists a rational mapping of $W \times W$ into W which maps $(c(\sigma), c(\tau))$ onto $c(\sigma\tau)$. It is now a simple matter to see that this rational mapping is a normal law of composition on W. Therefore, by the Weil-Rosenlicht result just mentioned, there exists a group variety G, and a birational mapping g' between W and G, such that $g'(c(\sigma\tau)) = g'(c(\sigma))g'(c(\tau))$ whenever σ, τ are independent generic elements of $\mathfrak{G}$. If for each generic element σ of $\mathfrak{G}$ we define $g(\sigma) = g'(c(\sigma))$ then g is a one-to-one mapping of the set of generic elements of $\mathfrak{G}$ onto the set of generic points of G, such that $\mathfrak{k}(\sigma) = \mathfrak{k}(g(\sigma))$ for every generic element σ of $\mathfrak{G}$, and such that $g(\sigma\tau) = g(\sigma)g(\tau)$ for all pairs σ, τ of independent generic elements of $\mathfrak{G}$; it is easy to see, moreover, that if σ, τ, σ', τ' are generic elements of $\mathfrak{G}$ then (σ', τ') is a specialization of (σ, τ) if and only if $(g(\sigma'), g(\tau'))$ is a specialization of $(g(\sigma), g(\tau))$. Every generic element σ of $\mathfrak{G}$ thus has the property that $g(\sigma\tau)g(\tau)^{-1}$ has the same value for all choices of generic element τ of $\mathfrak{G}$ such that σ, τ are independent. It is now a simple matter to prove that every element σ of $\mathfrak{G}$ has this property. Indeed, if $\sigma \in \mathfrak{G}$ and if τ, τ' are generic elements of $\mathfrak{G}$ such that σ, τ are independent and σ, τ' are independent, we may find a generic element τ'' of $\mathfrak{G}$ such that σ, τ, τ'' are independent and σ, τ', τ'' are independent; obviously then $\sigma\tau''$, $\tau''^{-1}\tau$ are independent generic points of $\mathfrak{G}$ as are $\tau'', \tau''^{-1}\tau$, so that

$$g(\sigma\tau)g(\tau)^{-1} = g(\sigma\tau'' \cdot \tau''^{-1}\tau)g(\tau'' \cdot \tau''^{-1}\tau)^{-1}$$
$$= g(\sigma\tau'')g(\tau''^{-1}\tau)\,(g(\tau'')g(\tau''^{-1}\tau))^{-1} = g(\sigma\tau'')g(\tau'')^{-1},$$

and similarly $g(\sigma\tau')g(\tau')^{-1} = g(\sigma\tau'')g(\tau'')^{-1}$, whence $g(\sigma\tau)g(\tau)^{-1} = g(\sigma\tau')g(\tau')^{-1}$.

This being established, we may define a mapping f of $\mathfrak{G}$ into G by the formula $f(\sigma) = g(\sigma\tau)g(\tau)^{-1}$ $(\sigma \in \mathfrak{G})$, where τ represents any generic element of $\mathfrak{G}$ such that σ, τ are independent. If $\sigma, \sigma' \in \mathfrak{G}$ we may choose a generic element τ such that each of the three pairs σ, τ and σ', τ and $\sigma, \sigma'\tau$ is independent, so that

$$f(\sigma\sigma') = g(\sigma\sigma'\tau)g(\tau)^{-1} = g(\sigma \cdot \sigma'\tau)g(\sigma'\tau)^{-1} \cdot g(\sigma'\tau)g(\tau)^{-1} = f(\sigma)f(\sigma')$$

and f is a homomorphism of the group $\mathfrak{G}$ into the group G. If ν is any element of the kernel of f and τ is a generic element of $\mathfrak{G}$ such that ν, τ are independent, so that $\nu\tau$ is generic on $\mathfrak{G}$, then $g(\nu\tau) = f(\nu\tau) = f(\tau) = g(\tau)$, whence $\nu\tau = \tau$ and ν is the identity; thus, the kernel of f is trivial. For any $\sigma \in \mathfrak{G}$ we may find a generic element τ of $\mathfrak{G}$ such that $\mathfrak{k}$ $(\mathfrak{k}(\sigma), \mathfrak{k}(f(\sigma)))$ and $\mathfrak{k}(\tau)$ are linearly disjoint over $\mathfrak{k}$, whence in particular σ, τ are independent; for such τ we have

$$\mathfrak{k}(f(\sigma), g(\tau)) = \mathfrak{k}(g(\sigma\tau)g(\tau)^{-1}, g(\tau)) = \mathfrak{k}(g(\sigma\tau), g(\tau)) = \mathfrak{k}(\sigma\tau, \tau) = \mathfrak{k}(\sigma, \tau),$$

whence it easily follows that $\mathfrak{k}(f(\sigma)) = \mathfrak{k}(\sigma)$. Again, if σ and σ' are any element of $\mathfrak{G}$ and we choose generic elements τ and τ' of $\mathfrak{G}$ such that each pair σ, τ and σ', τ' is independent then ([7], Ch. II, Props. 6, 8, pp. 777, 778) σ' is a specialization of σ if and only if $(\tau'\sigma', \tau')$ is a specialization of $(\tau\sigma, \tau)$; this takes place if and only if $(g(\tau'\sigma'), g(\tau'))$ is a specialization of $(g(\tau\sigma), g(\tau))$, that is, if and only if $(f(\tau')f(\sigma'), f(\tau'))$ is a specialization of $(f(\tau)f(\sigma), f(\tau))$, which in turn occurs if and only if $f(\sigma')$ is a specialization of $f(\sigma)$. To prove that f is a birational isomorphism of $\mathfrak{G}$ onto G it remains to show that G is the image of f. Now, the image contains $f(\tau)$, where τ is a generic element of $\mathfrak{G}$, and $f(\tau)$ is a generic element of G; therefore the smallest subvariety of G conaining the image is G itself. It follows (with the help of [7], Ch. II, Prop. 9, p. 779) that the complement of the image in G is contained in a proper subvariety of G, so that (Ch. I, Prop. 3) the image of f is G.

2. Irreducible group variety onto Galois group. We now prove the following partial converse to Theorem 1.

THEOREM 2. *Every irreducible group variety defined over the algebraically closed field $\mathfrak{k}$ of characteristic* 0 *is birationally isomorphic to the Galois group of a strongly normal extension of a differential field with field of constants $\mathfrak{k}$.*

Proof. Let G be an irreducible group variety defined over $\mathfrak{k}$, and set $r = \dim G$. The field $\mathfrak{K}_0^*$ of all rational functions on G (not necessarily defined over $\mathfrak{k}$) is a finitely generated field extension of transcendence degree r of its subfield of all constant functions on G, which subfield we identify with $\mathfrak{k}^*$ by identifying each constant function with its value. The subfield $\mathfrak{K}_0$ of $\mathfrak{K}_0^*$ consisting of all rational functions on G for which $\mathfrak{k}$ is a field of definition, is a finitely generated field extension of transcendence degree r of $\mathfrak{k}$. $\mathfrak{K}_0$ and $\mathfrak{k}^*$ are linearly disjoint over $\mathfrak{k}$ (so that their intersection is $\mathfrak{k}$), and their compositum is $\mathfrak{K}_0^*$.

For each point $s \in G$ there exists a unique automorphism ρ_s of the field $\mathfrak{K}_0^*$ with the following property: if $f \in \mathfrak{K}_0^*$ and $t \in G$ then $\rho_s f$ is defined at t if and only if f is defined at ts and, when this is the case, $(\rho_s f)(t) = f(ts)$. The computation

$$((\rho_s\rho_{s'})f)(t) = (\rho_s(\rho_{s'}f))(t) = (\rho_{s'}f)(ts) = f(tss') = (\rho_{ss'}f)(t)$$

shows that

$$\rho_{ss'} = \rho_s\rho_{s'} \quad (s, s' \in G). \tag{1}$$

Also, if $\rho_s f = f$ for every $f \in \mathfrak{K}_0^*$, that is if $f(ts) = f(t)$ for all $f \in \mathfrak{K}_0^*$ and

a generic point t of G such that t, s are independent (over $\mathfrak{k}$), then obviously s is the unity element of G, that is

(2) "$\rho_s =$ identity automorphism" implies "$s =$ unity element of G."

If $s \in G$, and if t denotes any point of G (in particular, any generic point of G such that s, t are independent over $\mathfrak{k}$), then $\mathfrak{k}(t, ts) = \mathfrak{k}(t, s) = \mathfrak{k}(ts, s)$. It follows that

$$\mathfrak{H}_0(\rho_s \mathfrak{H}_0) = \mathfrak{H}_0(\mathfrak{k}(s)) = (\rho_s \mathfrak{H}_0)(\mathfrak{k}(s)). \tag{3}$$

Also if $s' \in G$ and we take t generic with s, t independent and s', t independent then s' is a specialization of s over $\mathfrak{k}$ if and only if s' is a specialization of s over $\mathfrak{k}(t)$, which occurs if and only if $(f(ts'))_{f \in \mathfrak{H}_0}$ is a specialization of of $(f(ts))_{f \in \mathfrak{H}_0}$ over $\mathfrak{k}(t)$, that is, $((\rho_{s'} f)(t))_{f \in \mathfrak{H}_0}$ is a specialization of $((\rho_s f)(t))_{f \in \mathfrak{H}_0}$ over $\mathfrak{k}(t)$; it follows that

$$\begin{cases} s' \text{ is a specialization of } s \text{ over } \mathfrak{k} \text{ if and only if} \\ (\rho_{s'} f)_{f \in \mathfrak{H}_0} \text{ is a specialization of } (\rho_s f)_{f \in \mathfrak{H}_0} \text{ over } \mathfrak{H}_0. \end{cases} \tag{4}$$

The set of all derivations of $\mathfrak{H}_0^*$ over $\mathfrak{k}^*$ is a vector space over $\mathfrak{H}_0^*$ of dimension r. It is known (for example, see Nakano [9]) that this space has a base $D_1, \cdots, D_r$ composed of derivations each of which maps $\mathfrak{H}_0$ into itself and is "right-invariant" in the sense that

$$\rho_s \circ D_i = D_i \circ \rho_s \qquad (s \in G, 1 \leqq i \leqq r). \tag{5}$$

An element of $\mathfrak{H}_0^*$ annihilated by $D_1, \cdots, D_r$ is annihilated by every derivation of $\mathfrak{H}_0^*$ over $\mathfrak{k}^*$, and therefore belongs to $\mathfrak{k}^*$.

Now let $(u_{ij})_{1 \leqq i \leqq r,\, 0 \leqq j < \infty}$ be a family of indeterminates, and let

$$\mathcal{E} = \mathfrak{k}((u_{ij})_{1 \leqq i \leqq r,\, 0 \leqq j < \infty}), \quad \mathfrak{H} = \mathfrak{H}_0((u_{ij})_{1 \leqq i \leqq r,\, 0 \leqq j < \infty});$$

the fields $\mathfrak{k}^*((u_{ij})_{1 \leqq i \leqq r,\, 0 \leqq j < \infty})$ and $\mathfrak{H}_0^*((u_{ij})_{1 \leqq i \leqq r,\, 0 \leqq j < \infty})$ may then be denoted by $\mathcal{E}(\mathfrak{k}^*)$ and $\mathfrak{H}(\mathfrak{k}^*)$, respectively. For each $s \in G$ the automorphism ρ_s of $\mathfrak{H}_0^*$ over $\mathfrak{k}^*$, can be extended to a unique automorphism (which we also denote by ρ_s) of $\mathfrak{H}(\mathfrak{k}^*)$ over $\mathcal{E}(\mathfrak{k}^*)$. For each i the invariant derivation D_i of $\mathfrak{H}_0^*$ over $\mathfrak{k}^*$ can be extended to a unique derivation (which we also denote by D_i) of $\mathfrak{H}(\mathfrak{k}^*)$ over $\mathfrak{k}^*$ with the property that $D_i u_{ij} = u_{i0}^{-1} u_{i,j+1}$ and $D_i u_{i'j} = 0$ whenever $i' \neq i$ $(0 \leqq j < \infty)$. For the extended automorphisms and derivations (1), (2), (5) continue to hold, and (3), (4) yield

$$\mathfrak{H}(\rho_s \mathfrak{H}) = \mathfrak{H}(\mathfrak{k}(s)) = (\rho_s \mathfrak{H})(\mathfrak{k}(s)) \qquad (s \in G) \tag{6}$$

$$\begin{cases} s' \text{ is a specialization of } s \text{ over } \mathfrak{k} \text{ if and only if} \\ (\rho_{s'} \alpha)_{\alpha \in \mathfrak{H}} \text{ is a specialization of } (\rho_s \alpha)_{\alpha \in \mathfrak{H}} \text{ over } \mathfrak{H}. \end{cases} \tag{7}$$

Furthermore, if α is any element of $\mathscr{H}(\mathscr{L}^*)$, then α may be written as a rational fraction with coefficients in $\mathscr{H}_0{}^*$ in the indeterminate u_{ij}, and if α contains u_{kl} effectively but for every $l' > l$ fails to contain $u_{kl'}$ effective then it is clear that $u_{k,l+1}$ is effectively present in $D_k\alpha$, so that $D_k\alpha \neq 0$; thus the only elements of $\mathscr{H}(\mathscr{L}^*)$ annihilated by every D_i are the elements of $\mathscr{L}^*$:

$$D_i\alpha = 0 (1 \leqq i \leqq r) \text{ if and only if } \alpha \in \mathscr{L}^*. \tag{8}$$

We now make $\mathscr{H}(\mathscr{L}^*)$ into a differential field (not necessarily contained in the universal extension $\mathscr{F}^*$ of $\mathscr{F}$ introduced in Section 1 of Chapter I) with derivation operators $\delta_1, \cdots, \delta_m$ by defining

$$\delta_1\alpha = \sum u_{i0}D_i\alpha, \qquad \delta_k\alpha = 0 (2 \leqq k \leqq m) \tag{9}$$

for every $\alpha \in \mathscr{H}(\mathscr{L}^*)$. It is clear that $\mathscr{L}$, $\mathscr{E}$, $\mathscr{H}$, $\mathscr{L}^*$ all are differential subfields of $\mathscr{H}(\mathscr{L}^*)$, so that $\mathscr{H}(\mathscr{L}^*)$ may now be written as $\mathscr{H}\langle\mathscr{L}^*\rangle$; furthermore, if we denote u_{i0} by u_i then $u_{ij} = \delta_1{}^j u$ $(1 \leqq i \leqq r, 0 \leqq j < \infty)$, so that we may write $\mathscr{E} = \mathscr{L}\langle u_1, \cdots, u_r\rangle$ and $\mathscr{H} = \mathscr{E}\langle\mathscr{H}_0\rangle$.

If $\alpha \in \mathscr{H}\langle\mathscr{L}^*\rangle$ contains u_{kl} effectively but for every $l' > l$ fails to contain $u_{kl'}$ effectively, then $\delta_1\alpha = \sum u_{i0}D_i\alpha$ contains $u_{k,l+1}$ effectively, so that $\delta_1\alpha \neq 0$. Therefore if $\delta_i\alpha = 0$ then $\alpha \in \mathscr{H}_0{}^*$, and from the equation $\sum u_{i0}D_i\alpha = 0$ and the fact that $u_{10}, \cdots, u_{r0}$ are linearly independent over $\mathscr{H}_0{}^*$ we conclude that $D_i\alpha = 0$ $(1 \leqq i \leqq r)$. It follows from (8) that the field of constants of the differential field $\mathscr{H}\langle\mathscr{L}^*\rangle$ is $\mathscr{L}^*$; $\mathscr{L}^*$ is therefore also the field of constants of $\mathscr{E}\langle\mathscr{L}^*\rangle$, and $\mathscr{L}$ is the field of constants of $\mathscr{H}$ and of $\mathscr{E}$.

Before proceeding farther, we rid ourselves of the awkward circumstance that $\mathscr{H}\langle\mathscr{L}^*\rangle$ is not in general a differential subfield of $\mathscr{F}^*$. In the first place, $\mathscr{H} = \mathscr{E}\langle\mathscr{H}_0\rangle = \mathscr{L}\langle u_1, \cdots, u_r, \mathscr{H}_0\rangle$, so that $\mathscr{H}$ is a finitely generated differential field extension of $\mathscr{L}$; since $\mathscr{F}^*$ is a universal extension of $\mathscr{L}$ (see [7], Ch. I, § 5, p. 768), there exists an isomorphism over $\mathscr{L}$ of $\mathscr{H}$ onto a differential subfield $\mathscr{H}'$ of $\mathscr{F}^*$. Secondly, a family of elements of $\mathscr{H}$ linearly independent over $\mathscr{L}$ is also linearly independent over $\mathscr{L}^*$ (see [7], Ch. I, § 4, Prop. 3 and remark following, pp. 766-767), so that $\mathscr{H}$ and $\mathscr{L}^*$ are linearly disjoint over $\mathscr{L}$; it follows that the isomorphism over $\mathscr{L}$ of $\mathscr{H}$ onto $\mathscr{H}'$ can be extended to a unique isomorphism over $\mathscr{L}^*$ of $\mathscr{H}\langle\mathscr{L}^*\rangle$ onto $\mathscr{H}'\langle\mathscr{L}^*\rangle$. By virtue of this isomorphism we identify $\mathscr{H}\langle\mathscr{L}^*\rangle$ with $\mathscr{H}'\langle\mathscr{L}^*\rangle$, so that henceforth all our differential fields are differential subfields of $\mathscr{F}^*$.

By (5) and (9) we know that $\rho_s\delta_i\alpha = \delta_i\rho_s\alpha$ $(s \in G, 1 \leqq i \leqq m, \alpha \in \mathscr{H}\langle\mathscr{L}^*\rangle)$, so that each ρ_s is an automorphism of $\mathscr{H}\langle\mathscr{L}^*\rangle$ as a differential field (not merely as a field). This proves ([7], Ch. II, § 3, Prop. 4, p. 775) that each ρ_s, identified with its restriction to $\mathscr{H}$, is a strong isomorphism of $\mathscr{H}$ over $\mathscr{E}$.

It follows from (6) that if t is a generic point of G then the transcendence degree of $\mathfrak{H}\langle\rho_t\mathfrak{H}\rangle$ over $\mathfrak{H}$ equals r, which in turn equals the transcendence degree of $\mathfrak{H}$ over $\mathcal{E}$. Consequently ([7], Ch. II, § 2, Prop. 2, p, 773) ρ_t is an isolated isomorphism of $\mathfrak{H}$ over $\mathcal{E}$. Since $\mathcal{E}$ is obviously relatively algebraically closed in $\mathfrak{H}$, it follows ([7], Ch. II, § 2, Prop. 3, p. 774) that every isomorphism of $\mathfrak{H}$ over $\mathcal{E}$ is a specialization of ρ_t, and therefore ([7], Ch. II, § 4, Prop. 5, p. 776) is strong. Thus $\mathfrak{H}$ is a strongly normal extension of $\mathcal{E}$; we denote its Galois group by $\mathfrak{H}$.

By (1), the mapping $f: s \to \rho_s$ $(s \in G)$ is a homomorphism of G into $\mathfrak{H}$. By (2) the kernel of f is trivial. By (6) and [7], Ch. I, § 4, Cor. 5 to Prop. 3, p. 768,

$$\mathcal{C}(\rho_s) = \mathcal{C}(s) \qquad (s \in G),$$

and by (7), $\rho_{s'}$ is a specialization of ρ_s if and only if s' is a specialization of s (over $\mathcal{C}$). Therefore f is a rational homomorphism of G into $\mathfrak{H}$. Since the Galois group $\mathfrak{H}$ is by Theorem 1 birationally isomorphic with a group variety, it follows from the corollary to Proposition 5 in Chapter I that the image of f is an algebraic subgroup of $\mathfrak{H}$. Since the image contains a generic point $f(t)$ of $\mathfrak{H}$, it follows that the image of f is $\mathfrak{H}$ itself. Again making use of the fact that $\mathfrak{H}$ is birationally isomorphic with a group variety, we conclude from the corollary to Proposition 6 in Chapter I that f is a birational isomorphism of G onto $\mathfrak{H}$.

3. Applications. It follows from Theorem 1 that many propositions about group varieties can be generalized by permitting each group variety in the statement of a given proposition to be any algebraic group (that is, either a Galois group or a group variety). For example (as we have already had occasion to observe in Section 2):

PROPOSITION 1. *Proposition 5 of Chapter* I *and its corollary, and Proposition 6 of Chapter* I *and its corollary can be generalized to arbitrary algebraic groups.*

Another example:

PROPOSITION 2. *If* $\sigma_1, \cdots, \sigma_n \in \mathfrak{G}$ *and if* $(\sigma_1', \cdots, \sigma_n')$ *is a specialization of* $(\sigma_1, \cdots, \sigma_n)$, *then* $(W_i(\sigma'_1, \cdots, \sigma_n'))_{i \in I}$ *is a specialization of* $(W_i(\sigma_1, \cdots, \sigma_n))_{i \in I}$, *where* $(W_i(X_1, \cdots, X_n))_{i \in I}$ *is any family of words in* n *indeterminates* $X_1, \cdots, X_n$.

This is an immediated consequence of the fact that the same proposition holds for group varieties.

On the other hand some propositions on group varieties can not be indiscriminately generalized in this way because such a generalization may depend on Theorem 2, which we have proved only for *irreducible* group varieties. For example, we have not the right to assert the generalization of Proposition 7 of Chapter I, or of its corollary.

It is perhaps of interest to observe that not only can we prove results on Galois groups by appealing to previously proved analogous results on group varieties, but also we can work in the opposite direction. For example, we can prove that, roughly speaking, a factor group of a group variety is itself a group variety. More precisely:

PROPOSITION 3. *If G is a group variety and K is a normal algebraic subgroup of G then, there exists a rational homomorphism of G with kernel K and image a group variety.*

This result is contained (except for the rather trivial extension here to group varieties which are not necessarily irreducible) in much stronger results of Nakano [10]; the novelty here lies in the fact that the result on group varieties is obtained as a consequence of theorems on algebraic differential equations. The proof runs as follows. By Proposition 1 of Chapter I, the result holds for a group variety G and kernel K provided it holds for the group variety G^0 (= the component of the unity of G) and kernel $G^0 \cap K$, that is, we may assume that G is irreducible. Making this assumption, we know from Theorem 2 that G is birationally isomorphic with a Galois group, say f is a birational isomorphism of G onto $\mathfrak{G}$; $f(K)$ is a *normal* algebraic subgroup of $\mathfrak{G}$ and therefore is the Galois group $\mathfrak{G}(\mathfrak{H})$ of $\mathfrak{L}$ over a differential field $\mathfrak{H}$ intermediate to $\mathfrak{F}$ and $\mathfrak{L}$ and *strongly normal* over $\mathfrak{F}$. Therefore (see Example 1 of Chapter I, § 4) the mapping r, which to each $\sigma \in \mathfrak{G}$ corresponds the restriction of σ to $\mathfrak{H}$, is a rational homomorphism of $\mathfrak{G}$ with kernel $\mathfrak{G}(\mathfrak{H})$ and image the Galois group $\mathfrak{H}$ of $\mathfrak{H}$ over $\mathfrak{F}$. By Theorem 1, there exists a birational isomorhpism h of $\mathfrak{H}$ onto a group variety H. It is now immediate that the composite mapping $h \circ r \circ f$ is a rational homomorphism of G with kernel K and image H.

Similarly, we can prove the following result.

PROPOSITION 4. *If a rational homomorphism of an algebraic group G has kernel K and image H, then* $\dim G = \dim K + \dim H$.

This too is contained in results of Nakano [10]. Our proof uses the fact that the dimension of the Galois group of a strongly normal differential field extension equals the transcendence degree of the extension. We omit the details, which are rather obvious.

Finally, we generalize to arbitrary connected group varieties (but only in the case of field characteristic 0, of course) a theorem on algebraic matric groups (Chevalley [3], Ch. II, § 4, Th. 1, part 2), p. 84, or Chevalley and Kolchin [4], § 1, Th. 2).

PROPOSITION 5. *The notation being as in § 2, if W is a subset of an irreducible group variety G then a necessary and sufficient condition for W to be an algebraic subgroup of G is that there exist a finite subset S, of the field $\mathfrak{H}_0$ of rational functions on G which are defined over $\mathfrak{k}$, such that W is the set of all points $s \in G$ for which $\rho_s \alpha = \alpha$ for every $\alpha \in S$.*

Indeed, we saw in the proof of Theorem 2 that $\mathfrak{H} = \mathfrak{E}\langle \mathfrak{H}_0 \rangle$ is a strongly normal extension of $\mathfrak{E} = \mathfrak{k}\langle u_1, \cdots, u_r \rangle$, and $s \to \rho_s$ $(s \in G)$ is a birational isomorphism f of G onto the Galois group $\mathfrak{H}$ of $\mathfrak{H}$ over $\mathfrak{E}$. W is an algebraic subgroup of G if and only if $f(W)$ is an algebraic subgroup of $\mathfrak{H}$, that is, is the Galois group of $\mathfrak{H}$ over a differential field $\mathfrak{E}_1$ intermediate to $\mathfrak{E}$ and $\mathfrak{H}$; now every intermediate differential field is a finitely generated differential field extension, and it is easy to see that $\mathfrak{E}_1 = \mathfrak{E}\langle \alpha_1, \cdots, \alpha_p \rangle$ for some finite set S of elements $\alpha_1, \cdots, \alpha_p$ which belong to $\mathfrak{H}_0$. Therefore $f(W)$ is the Galois group of $\mathfrak{H}$ over $\mathfrak{E}_1$ if and only if W is the set of all elements $s \in G$ such that $\rho_s \alpha_i = \alpha_i$ $(1 \leqq i \leqq p)$.

Chapter III. One-dimensional cohomology.

1. Definition of $H^1(\mathfrak{G}, G)$. The universal extension $\mathfrak{F}^*$ of $\mathfrak{F}$ is, like its field of constants $\mathfrak{k}^*$, an algebraically closed field extension of $\mathfrak{k}$ of infinite transcendence degree and therefore may be used, like $\mathfrak{k}^*$, as a universal domain for algebraic geometry. Furthermore, since $\mathfrak{k}^* \subset \mathfrak{F}^*$, every variety V in the algebraic geometry based on $\mathfrak{k}^*$ as universal domain determines in an obvious way a variety $V^\dagger$ in the algebraic geometry based on $\mathfrak{F}^*$ as universal domain, such that a point of $V^\dagger$ belongs to V if and only if the point is rational over $\mathfrak{k}^*$. It is obvious that if V is a group variety in the algebraic geometry based on $\mathfrak{k}^*$ then $V^\dagger$ is a group variety in the algebraic geometry based on $\mathfrak{F}^*$.

Now, a mapping f of the Galois group $\mathfrak{G}$ of $\mathfrak{H}$ over $\mathfrak{F}$ into a group variety G may also be regarded as a mapping of $\mathfrak{G}$ into $G^\dagger$; since $\mathfrak{k}(f(\sigma))$ consists of constants for any $\sigma \in \mathfrak{G}$, the condition "$\mathfrak{k}(f(\sigma)) \subset \mathfrak{k}(\sigma)$" is equivalent to "$\mathfrak{H}(f(\sigma)) \subset \mathfrak{H}(\sigma\mathfrak{H})$," for any $\sigma, \sigma' \in \mathfrak{G}$ the condition "$f(\sigma')$ is a specialization of $f(\sigma)$ over $\mathfrak{k}$" is equivalent to the condition "$f(\sigma')$ is a specialization of $f(\sigma)$ over $\mathfrak{H}$" and the condition "$f(\sigma\sigma') = f(\sigma)f(\sigma')$" may, if we wish, be written in the form "$f(\sigma\sigma') = f(\sigma) \cdot \sigma f(\sigma')$." Therefore the following definition generalizes the notion of rational homomorphism of $\mathfrak{G}$ into G:

By a *rational crossed homomorphism*, or a *cocycle of dimension* 1, or a 1-*cocycle*, of the Galois group $\mathfrak{G}$ into the group variety G we mean any mapping f of $\mathfrak{G}$ *into* $G^\dagger$ such that $\mathfrak{F}(f(\sigma)) \subset \mathfrak{F}(\sigma\mathfrak{F})$ for all $\sigma \in \mathfrak{G}$, such that "σ' is a specialization of σ" implies "$f(\sigma')$ is a specialization of $f(\sigma)$ over $\mathfrak{F}$" for all $\sigma, \sigma' \in \mathfrak{G}$, and such that $f(\sigma\sigma') = f(\sigma) \cdot \sigma f(\sigma')$ for all $\sigma, \sigma' \in \mathfrak{G}$.

As remarked above, every rational homomorphism of $\mathfrak{G}$ into G is a rational crossed homomorphism of $\mathfrak{G}$ into G. Also, if α is any point of $G^\dagger$ *which is rational over* $\mathfrak{F}$, then it is evident that the mapping $\sigma \to \alpha^{-1}\sigma\alpha$ $(\sigma \in \mathfrak{G})$ is a rational crossed homomorphism of $\mathfrak{G}$ into G; such a rational crossed homomorphism is said to be *trivial*, or to be a *coboundary of dimension* 1 or a 1-*coboundary* of $\mathfrak{G}$ into G.

We denote the set of all 1-cocyles of $\mathfrak{G}$ into G by $Z^1(\mathfrak{G}, G)$, and the set of all 1-coboundaries by $B^1(\mathfrak{G}, G)$, so that $B^1(\mathfrak{G}, G) \subset Z^1(\mathfrak{G}, G)$. The mapping which to each element of $\mathfrak{G}$ corresponds the unity element 1 of G is obviously a 1-coboundary (because $1^{-1}\sigma 1 = 1$ for every $\sigma \in \mathfrak{G}$); we denote it by 1. In general, as we have made no assumption of commutativity for G, $Z^1(\mathfrak{G}, G)$ and $B^1(\mathfrak{G}, G)$ are not groups under the law of composition $(f, g) \to fg$, where $(fg)(\sigma) = f(\sigma)g(\sigma)$, because fg need not be an element of $Z^1(\mathfrak{G}, G)$.

If $f \in Z^1(\mathfrak{G}, G)$, and if $\alpha \in G^\dagger$ is rational over $\mathfrak{F}$, then the mapping g of $\mathfrak{G}$ defined by $g(\sigma) = \alpha^{-1}f(\sigma)\sigma\alpha$ is obviously an element of $Z^1(\mathfrak{G}, G)$, which we call *cohomologous* to f. The relation "g is cohomologous to f" is an equivalence on $Z^1(\mathfrak{G}, G)$. An equivalence class with respect to this equivalence we call a *cohomology class of dimension* 1 or a 1-*cohomology class* of $\mathfrak{G}$ into G; we denote the set of all these by $H^1(\mathfrak{G}, G)$. $B^1(\mathfrak{G}, G)$ itself is a 1-cohomology class; when we think of this class as an element of $H^1(\mathfrak{G}, G)$ we frequently permit ourselves to denote it by 1, and we also permit ourselves to write $H^1(\mathfrak{G}, G) = 1$ when we wish to assert that $B^1(\mathfrak{G}, G)$ is the only 1-cohomology class of $\mathfrak{G}$ into G.

Of course, if G is commutative then $Z^1(\mathfrak{G}, G)$ and $B^1(\mathfrak{G}, G)$ are commutative groups, and $H^1(\mathfrak{G}, G)$ is the factor group $Z^1(\mathfrak{G}, G)/B^1(\mathfrak{G}, G)$; we then call it the *first cohomology group* of $\mathfrak{G}$ into G.

2. Determination of $H^1(\mathfrak{G}, G)$ in some general cases. The computation of $H^1(\mathfrak{G}, G)$ in various cases below is facilitated by the observation that two 1-cocycles f and g of $\mathfrak{G}$ into G are equal provided $f(\sigma) = g(\sigma)$ for all $\sigma \in \mathfrak{G}$ which are automorphisms of $\mathfrak{F}$ over $\mathfrak{F}$); this follows from [7]. Ch. II, Cor. 2 to Prop. 9, p. 783.

Our results on $H^1(\mathfrak{G}, G)$ are summarized in the following theorem. We recall that $\mathfrak{G}$ is the Galois group of the strongly normal extension $\mathfrak{F}$

of the differential field $\mathfrak{F}$ of characteristic 0 with algebraically closed field of constants $\mathfrak{C}$. G denotes any group variety defined over $\mathfrak{C}$. We denote the group of all invertible square matrices of degree n with constant coordinates by $GL(n)$; $GL(n)$ is an irreducible group variety of dimension n^2.

THEOREM 1. 1) $H^1(\mathfrak{G}, GL(n)) = 1$.

2) *If* $\mathfrak{F}$ *is algebraically closed then* $H^1(\mathfrak{G}, G) = 1$.

3) *If* G *is commutative then every element of the group* $H^1(\mathfrak{G}, G)$ *is of finite order.*

Proof of 1). Let $f \in Z^1(\mathfrak{G}, GL(n))$, and let τ be a generic element of the component of the identity $\mathfrak{G}^0$ of $\mathfrak{G}$. Since $\mathfrak{H}(f(\tau)) \subset \mathfrak{H}(\tau\mathfrak{H})$, there must exist a finite number of elements $\tau\xi_1, \cdots, \tau\xi_p \in \tau\mathfrak{H}$ which are algebraically independent over $\mathfrak{H}$, and a finite number of elements $\tau\eta_1, \cdots, \tau\eta_q \in \tau\mathfrak{H}$ which are linearly independent over $\mathfrak{H}(\tau\xi_1, \cdots, \tau\xi_p)$, such that

$$\mathfrak{H}(f(\tau)) \subset \sum_{j=1}^{q} \mathfrak{H}(\tau\xi_1, \cdots, \tau\xi_p) \cdot \tau\eta_j.$$

Therefore there exist matrices $A_j = A_j(X_1, \cdots, X_p)$, $1 \leqq j \leqq q$, square and of degree n, with coordinates which are polynomials over $\mathfrak{H}$ in indeterminates $X_1, \cdots, X_p$, and there exists a nonzero polynomial

$$B = B(X_1, \cdots, X_p) \in \mathfrak{H}[X_1, \cdots, X_p],$$

such that

$$f(\tau) = \sum_{j=1}^{q} (\tau\eta_j) B(\tau\xi_1, \cdots, \tau\xi_p)^{-1} A_j(\tau\xi_1, \cdots, \tau\xi_p). \tag{1}$$

This equation continues to hold, of course, if τ is replaced by any generic element of $\mathfrak{G}^0$. We assume without loss of generality that B does not have a factor of degree > 0 which divides every coordinate of every A_j, and we assume, too, that one of the coefficients in B is 1. For any $\sigma \in \mathfrak{G}^0$ which is an *auto*morphism of $\mathfrak{H}$, $\sigma^{-1}\tau$ is a generic element of $\mathfrak{G}^0$. On the other hand, $f(\tau) = f(\sigma \cdot \sigma^{-1}\tau) = f(\sigma) \cdot \sigma f(\sigma^{-1}\tau)$. Therefore by (1)

$$\sum_{j=1}^{q} (\tau\eta_j) B(\tau\xi_1, \cdots, \tau\xi_p)^{-1} A_j(\tau\xi_1, \cdots, \tau\xi_p)$$
$$= f(\sigma) \cdot \sum_{j=1}^{q} (\tau\eta_j) B_\sigma(\tau\xi_1, \cdots, \tau\xi_p)^{-1} A_{j\sigma}(\tau\xi_1, \cdots, \tau\xi_p),$$

where B_σ is obtained from B by replacing each coefficient by its image under σ, and similarly for $A_{j\sigma}$. Because $\tau\eta_1, \cdots, \tau\eta_q$ are linearly independent over $\mathfrak{H}(\tau\xi_1, \cdots, \tau\xi_p)$, it follows that

$$B(\tau\xi_1, \cdots, \tau\xi_p)^{-1} A_j(\tau\xi_1, \cdots, \tau\xi_p)$$
$$= f(\sigma) B_\sigma(\tau\xi_1, \cdots, \tau\xi_p)^{-1} A_{j\sigma}(\tau\xi_1, \cdots, \tau\xi_p), \qquad 1 \leqq j \leqq p.$$

Because $\tau\xi_1, \cdots, \tau\xi_p$ are algebraically independent over $\mathfrak{H}$, it follows that

$$B_\sigma A_j = f(\sigma) \cdot BA_{j\sigma}, \qquad 1 \leqq j \leqq p.$$

Because no factor of B of degree > 0 divides every coordinate of every A_j, it follows that B_σ is a multiple of B by an element of $\mathfrak{H}$, and because a coefficient in B is 1 we conclude that $B_\sigma = B$, so that

$$A_j = f(\sigma) A_{j\sigma}, \qquad 1 \leqq j \leqq p. \tag{2}$$

Now, it is a consequence of (1) that if $Y_1, \cdots, Y_q$ are additional indeterminates then

$$\det\left(\sum_{j=1}^{q} Y_j A_j(X_1, \cdots, X_p)\right) \neq 0.$$

Therefore there exist element $a_1, \cdots, a_p, b_1, \cdots, b_q \in \mathfrak{F}$ such that if we form the matrix $\alpha = \sum b_j A_j(a_1, \cdots, a_p)$ then $\det \alpha \neq 0$; obviously the coordinates of α are elements of $\mathfrak{H}$. However,

$$\sigma A_j(a_1, \cdots, a_p) = A_{j\sigma}(\sigma a_1, \cdots, \sigma a_p) = A_{j\sigma}(a_1, \cdots, a_p).$$

Therefore by (2) we have $\alpha = f(\sigma)\sigma\alpha$. Thus, if we let g be the 1-cocycle cohomologous to f defined by the equation $g(\sigma') = \alpha^{-1} f(\sigma')\sigma'\alpha$ for all $\sigma' \in \mathfrak{G}$, then $g(\sigma) = 1$ for every $\sigma \in \mathfrak{G}^0$.

If σ' and σ'' belong to the same coset of $\mathfrak{G}$ relative to $\mathfrak{G}^0$, then $g(\sigma'') = g(\sigma' \cdot \sigma'^{-1}\sigma'') = g(\sigma')\sigma' g(\sigma'^{-1}\sigma'') = g(\sigma')$, so that g is constant on each coset. Since every coset contains an automorphism of $\mathfrak{H}$, $g(\sigma')$ is always rational over $\mathfrak{H}$; but, for any $\sigma \in \mathfrak{G}^0$, we have $g(\sigma') = g(\sigma\sigma') = g(\sigma)\sigma g(\sigma') = \sigma g(\sigma')$, so that $g(\sigma')$, is always rational over the relative algebraic closure $\mathfrak{F}^0$ of $\mathfrak{F}$ in $\mathfrak{H}$. Of course, if $\sigma_1, \cdots, \sigma_h$ form a complete set of representatives of the cosets of $\mathfrak{G}$ relative to $\mathfrak{G}^0$, then the restrictions of $\sigma_1, \cdots, \sigma_h$ to $\mathfrak{F}^0$ are precisely the elements of the Galois group of $\mathfrak{F}^0$ over $\mathfrak{F}$. Now, if $T_1, \cdots, T_h$ are indeterminates, then $\det(\sum T_i g(\sigma_i))$ is a polynomial in $\mathfrak{F}^0[T_1, \cdots, T_h]$, and is different from 0 (because, for example, at $(1, 0, \cdots, 0)$ it has the value $\det(g(\sigma_i))$). Therefore (Bourbaki [1], Ch. V, § 10, Th. 4, p. 157) there exists an element $a \in \mathfrak{F}^0$ such that $\det(\sum(\sigma_i a) g(\sigma_i)) \neq 0$. Writing $\beta = \sum(\sigma_i a) g(\sigma_i)$, we have for any $\sigma' \in \mathfrak{G}$,

$$\sigma'\beta = \sum(\sigma'\sigma_i a)\sigma' g(\sigma_i) = \sum(\sigma'\sigma_i a) g(\sigma')^{-1} g(\sigma'\sigma_i) = g(\sigma')^{-1}\beta.$$

Letting $\gamma = \beta^{-1}\alpha^{-1}$ we therefore obtain $\gamma^{-1}\sigma'\gamma = \alpha\beta\sigma'\beta^{-1} \cdot \sigma'\alpha^{-1} = \alpha g(\sigma')\sigma'\alpha^{-1} = f(\sigma')$ for every $\sigma' \in \mathfrak{G}$, so that $f \in B^1(\mathfrak{G}, GL(n))$.

Proof of 2). Since $\mathfrak{F}$ is algebraically closed $\mathfrak{G}$ is irreducible. Let $f \in Z^1(\mathfrak{G}, G)$, and let τ be a generic element of $\mathfrak{G}$. Letting $\eta_1, \cdots, \eta_n$ be elements of $\mathfrak{H}$ such that $\mathfrak{H} = \mathfrak{F}(\eta_1, \cdots, \eta_n)$, we see that

$$\mathfrak{H}(\tau\mathfrak{H}) = \mathfrak{F}(\eta_1, \cdots, \eta_n, \tau\eta_1, \cdots, \tau\eta_n);$$

also, $(\eta_1, \cdots, \eta_n)$ and $(\tau\eta_1, \cdots, \tau\eta_n)$ are independent over $\mathfrak{F}$ (because the transcendence degree of $\mathfrak{F}(\eta_1, \cdots, \eta_n, \tau\eta_1, \cdots, \tau\eta_n)$ over $\mathfrak{F}$ equals that of $\mathfrak{F}(\eta_1, \cdots, \eta_n)$ over $\mathfrak{F}$ plus that of $\mathfrak{F}(\tau\eta_1, \cdots, \tau\eta_n)$ over $\mathfrak{F}$).

Let $f_i(\tau)$ denote a representative in affine space of the abstract point $f(\tau)$. Because $f(\tau)$ is rational over $\mathfrak{G}(\tau\mathfrak{G})$ there exist rational fractions $R_{i1}, \cdots, R_{id_i}$ in $2n$ indeterminates, with coefficients in $\mathfrak{F}$ and denominators different from 0 at $(\eta_1, \cdots, \eta_n, \tau\eta_1, \cdots, \tau\eta_n)$, such that

$$f_i(\tau) = (R_{i1}(\eta_1, \cdots, \eta_n, \tau\eta_1, \cdots, \tau\eta_n), \cdots, R_{id_i}(\eta_1, \cdots, \eta_n, \tau\eta_1, \cdots, \tau\eta_n))$$

For any $\sigma \in \mathfrak{G}$ which is an *auto*morphism of $\mathfrak{G}$, the corresponding representative of $\sigma f(\sigma^{-1}\tau)$ is obviously given by

$$\sigma f_i(\sigma^{-1}\tau) = (R_{i1}(\sigma\eta_1, \cdots, \sigma\eta_n, \tau\eta_1, \cdots, \tau\eta_n), \cdots, R_{id_i}(\sigma\eta_1, \cdots, \sigma\eta_n, \tau\eta_1, \cdots, \tau\eta_n)).$$

Since $\mathfrak{F}$ is algebraically closed, there exists an affine point $(b_1, \cdots, b_n)$ which is a specialization in the sense of algebraic geometry of $(\tau\eta_1, \cdots, \tau\eta_n)$ over $\mathfrak{F}$, with each $b_i \in \mathfrak{F}$ and with the denominator of each R_{ij} different from 0 at $(\eta_1, \cdots, \eta_n, b_1, \cdots, b_n)$; application of the automorphism σ to $R_{ij}(\eta_1, \cdots, \eta_n, b_1, \cdots, b_n)$ shows that the denominator of each R_{ij} is different from 0 at $(\sigma\eta_1, \cdots, \sigma\eta_n, b_1, \cdots, b_n)$. Since $(\eta_1, \cdots, \eta_n)$ and $(\tau\eta_1, \cdots, \tau\eta_n)$ are independent over $\mathfrak{F}$, $(\eta_1, \cdots, \eta_n, b_1, \cdots, b_n)$ is a specialization of $(\eta_1, \cdots, \eta_n, \tau\eta_1, \cdots, \tau\eta_n)$ over $\mathfrak{F}$, that is, $(b_1, \cdots, b_n)$ is a specialization of $(\tau\eta_1, \cdots, \tau\eta_n)$ *over* $\mathfrak{G}$. Therefore if we set

$$g_i = (R_{i1}(\eta_1, \cdots, \eta_n, b_1, \cdots, b_n), \cdots, R_{id_i}(\eta_1, \cdots, \eta_n, b_1, \cdots, b_n))$$

then $(g_i, \sigma g_i)$ is a specialization of $(f_i(\tau), \sigma f(\sigma^{-1}\tau))$ over $\mathfrak{G}$ for every i. Accordingly, the affine points g_i are representatives of an abstract point g which is rational over $\mathfrak{G}$, and $(g, \sigma g)$ is a specialization of $(f(\tau), \sigma f(\sigma^{-1}\tau))$ over $\mathfrak{G}$. But $\tau = \sigma \cdot \sigma^{-1}\tau$, so that $f(\tau) = f(\sigma)\sigma f(\sigma^{-1}\tau)$, and $f(\sigma)$ is rational over $\mathfrak{G}$. Therefore $g = f(\sigma)\sigma g$ for all $\sigma \in \mathfrak{G}$ which are automorphisms of $\mathfrak{G}$, hence for all $\sigma \in \mathfrak{G}$; that is, $f \in B^1(\mathfrak{G}, G)$.

Proof of 3). Let $\mathfrak{F}'$ be the algebraic closure of $\mathfrak{F}$. $\mathfrak{F}'$ is a normal algebraic extension of $\mathfrak{F}$, and $\mathfrak{F}' \cap \mathfrak{G}$ is the relative algebraic closure $\mathfrak{F}^0$ of $\mathfrak{F}$ in $\mathfrak{G}$. Therefore every automorphism τ of $\mathfrak{F}'$ over $\mathfrak{F}^0$ can be extended to a unique automorphism (which we also denote by τ) of the compositum $\mathfrak{F}'\langle\mathfrak{G}\rangle = \mathfrak{G}\langle\mathfrak{F}'\rangle$ over $\mathfrak{G}$ (the group of automorphisms of $\mathfrak{F}'$ over $\mathfrak{F}^0$ thereby becoming identified with the group of automorphisms of $\mathfrak{G}\langle\mathfrak{F}'\rangle$ over $\mathfrak{G}$); also ([7], Ch. III, § 3, Th. 5, p. 799) $\mathfrak{G}\langle\mathfrak{F}'\rangle$ is a strongly normal extension of $\mathfrak{F}'$, and every isomorphism σ of $\mathfrak{G}$ over $\mathfrak{F}^0$ can be extended to a unique isomorphism (which we also denote by σ) of $\mathfrak{G}\langle\mathfrak{F}'\rangle$ over $\mathfrak{F}'$ (the component of identity $\mathfrak{G}^0$ of $\mathfrak{G}$ thereby becoming identified with the Galois group of $\mathfrak{G}\langle\mathfrak{F}'\rangle$ over $\mathfrak{F}'$). Since every element of $\mathfrak{G}\langle\mathfrak{F}'\rangle$ can be expressed as the

quotient of two elements of the form $\sum \psi_i \phi_i$ ($\psi_i \in \mathfrak{H}, \phi_i \in \mathfrak{F}'$), and since when σ is an *auto*morphism

$$\sigma\tau(\sum \psi_i\phi_i) = \sigma \sum \psi_i \tau\phi_i = \sum (\sigma\psi_i)\tau\phi_i = \tau \sum (\sigma\psi_i)\phi_i = \tau\sigma(\sum \psi_i\phi_i),$$

we see that $\sigma\tau = \tau\sigma$. It follows that if $\alpha_1, \cdots, \alpha_h$ are the conjugates of $\alpha \in \mathfrak{H}\langle\mathfrak{F}'\rangle$ over $\mathfrak{H}$ then $\sigma\alpha_1, \cdots, \sigma\alpha_h$ are the conjugates of $\sigma\alpha$ over $\mathfrak{H}$ (where σ is any element of $\mathfrak{G}^0$ which is an automorphism of $\mathfrak{H}$).

Now let $f \in Z^1(\mathfrak{G}, G)$. Since $\mathfrak{G}^0$ can be identified with the Galois group of $\mathfrak{H}\langle\mathfrak{F}'\rangle$ over $\mathfrak{F}'$, since (for $\sigma \in \mathfrak{G}^0$) $\mathfrak{H}\langle\mathfrak{F}'\rangle(f(\sigma)) \subset \mathfrak{H}\langle\mathfrak{F}'\rangle(\sigma(\mathfrak{H}\langle\mathfrak{F}'\rangle))$, and since (for $\sigma, \sigma' \in \mathfrak{G}^0$) "$\sigma'$ is a specialization of σ (as isomorphism of $\mathfrak{H}\langle\mathfrak{F}'\rangle$)" implies "$\sigma'$ is a specialization of σ (as isomorphisms of $\mathfrak{H}$)" which implies "$f(\sigma')$ is a specialization of $f(\sigma)$ over $\mathfrak{H}$" which (because $\mathfrak{H}\langle\mathfrak{L}^*\rangle$ and $\mathfrak{H}\langle\mathfrak{F}'\rangle$ are linearly disjoint over $\mathfrak{H}$) in turn implies "$f(\sigma')$ is a specialization of $f(\sigma)$ over $\mathfrak{H}\langle\mathfrak{F}'\rangle$," we see that the restriction of f to $\mathfrak{G}^0$ may be considered as a rational crossed homomorphism of the Galois group of $\mathfrak{H}\langle\mathfrak{F}'\rangle$ over $\mathfrak{F}'$ into G. By 2) above, therefore, there exists an $\alpha \in \mathfrak{G}^\dagger$ rational over $\mathfrak{H}\langle\mathfrak{F}'\rangle$, such that $f(\sigma) = \alpha^{-1}\sigma\alpha$ for every $\sigma \in \mathfrak{G}^0$ which is an automorphism of $\mathfrak{H}$. Denoting the conjugates of α over $\mathfrak{H}$ by $\alpha_1, \cdots, \alpha_h$ we see by the above (since $f(\sigma)$ is rational over $\mathfrak{H}$) that $f(\sigma) = \alpha_i^{-1}\sigma\alpha_i$ ($1 \leqq i \leqq h$). Since G is by assumption commutative, we theerfore have $(f(\sigma))^h = \beta^{-1}\sigma\beta$, where $\beta = \alpha_1 \cdots \alpha_h$ is rational over $\mathfrak{H}$. Therefore the 1-cocycle g cohomologous to f^h, defined by $g(\sigma) = \beta(f(\sigma))^h\sigma\beta^{-1}$ ($\sigma \in \mathfrak{G}$), has the value 1 at every element of $\mathfrak{G}^0$. It follows, as in the proof of 1) above, that g is constant on each coset and that $g(\sigma)$ is rational over $\mathfrak{F}^0$ for every $\sigma \in \mathfrak{G}$. Consequently, letting $\sigma_1, \cdots, \sigma_k$ denote a complete set of representatives of the cosets of $\mathfrak{G}$ relative to $\mathfrak{G}^0$, and setting $\gamma = g(\sigma_1) \cdots g(\sigma_k)$, we find that

$$\sigma\gamma = \prod \sigma g(\sigma_j) = \prod g(\sigma)^{-1} g(\sigma\sigma_j) = (g(\sigma))^{-k}\gamma$$

for every $\sigma \in \mathfrak{G}$, so that $(f(\sigma))^{hk} = \beta^{-k}\gamma\sigma(\gamma^{-1}\beta^k)$ for every $\sigma \in \mathfrak{G}$ and $f^{hk} \in B^1(\mathfrak{G}, G)$.

3. Applications. Theorem 1 of Chapter II asserts that there exists a birational isomorphism of the Galois group $\mathfrak{G}$ onto a group variety G. Now such a birational isomorphism, or more generally any rational homomorphism f of $\mathfrak{G}$ into G, is a 1-cocycle of $\mathfrak{G}$ into G. If we can prove that f is a 1-coboundary then there exists a point α of G (or, more precisely, of $G^\dagger$) which is rational over $\mathfrak{H}$ such that $f(\sigma) = \alpha^{-1}\sigma\alpha$, that is, such that $\sigma\alpha = \alpha f(\sigma)$ for every $\sigma \in \mathfrak{G}$. In the case in which f is actually an isomorphism, we see that α is invariant under σ only when σ is the identity isomorphism, so that $\mathfrak{H} = \mathfrak{F}\langle\alpha\rangle$. More generally, if f is a homomorphism with *finite* kernel then $\mathfrak{H}$ is a normal algebraic extension of $\mathfrak{F}\langle\alpha\rangle$ with Galois group isomorphic to

the kernel of f. These remarks are the basis of the use of Theorem 1 in the proofs of the following two theorems.

THEOREM 2. *A necessary and sufficient condition that the strongly normal extension $\mathcal{G}$ of $\mathcal{F}$ be a Picard-Vessiot extension is that its Galois group $\mathfrak{G}$ be birationally isomorphic to an algebraic group of matrices.*

Proof. If $\mathcal{G}$ is a Picard-Vessiot extension of $\mathcal{F}$ then, as we have already observed (Chapter I, § 3, Example 2), $\mathfrak{G}$ is birationally isomorphic to an algebraic matric group. Conversely, let f be a birational isomorphism of $\mathfrak{G}$ onto an algebraic subgroup of $GL(n)$. By Theorem 1, there exists an invertible square matrix $\alpha = (\alpha_{ij})_{1\leq i\leq n,\, 1\leq j\leq n}$ with coordinates $\alpha_{ij} \in \mathcal{G}$ such that $\sigma\alpha = \alpha f(\sigma)$ for every $\sigma \in \mathfrak{G}$, and $\mathcal{G} = \mathcal{F}\langle\alpha\rangle$. Therefore if $\eta_1, \cdots, \eta_p$ form a maximal subset of the set of elements α_{ij} linearly independent over $\mathcal{C}$ (and therefore over every field of constants) then for each $\sigma \in \mathfrak{G}$ there exists an invertible matrix $(c_{ij}(\sigma))_{1\leq i\leq p,\, 1\leq j\leq p}$ with constant coordinates $c_{ij}(\sigma)$, such that $\sigma\eta_j = \sum_{1\leq i\leq p} c_{ij}(\sigma)\eta_i$ $(1 \leqq j \leqq p)$. It is now clear that if $\theta_1, \cdots, \theta_p$ are arbitrary operators of the form $\delta_1^{i_1} \cdots \delta_m^{i_m}$ then the determinant

$$W_{\theta_1 \cdots \theta_p}(\eta_1, \cdots, \eta_p) = \det(\theta_i \eta_j)_{1\leq i\leq p,\, 1\leq j\leq p}$$

has the property that

$$\sigma W_{\theta_1 \cdots \theta_p}(\eta_1, \cdots, \eta_p) = \det(c_{ij}(\sigma)) \cdot W_{\theta_1 \cdots \theta_p}(\eta_1, \cdots, \eta_p),$$

so that the ratio of any two nonzero determinants of this type is invariant under $\mathfrak{G}$ and hence is an element of $\mathcal{F}$. Thus ([6], § 3, p. 599) $\mathcal{G}$ is a Picard-Vessiot extension of $\mathcal{F}$.

COROLLARY. *If $\mathcal{G}$ is a Picard-Vessiot extension of $\mathcal{F}$ and $\mathcal{H}$ is any differential field between $\mathcal{F}$ and $\mathcal{G}$ which is a strongly normal extension*[3] *of $\mathcal{F}$ then $\mathcal{H}$ is a Picard-Vessiot extension of $\mathcal{F}$.*

Proof. Since $\mathcal{G}$ is Picard-Vessiot over $\mathcal{F}$ there exists a birational isomorphism f of $\mathfrak{G}$ onto an algebraic matric group G. The Galois group $\mathfrak{G}(\mathcal{H})$ of $\mathcal{G}$ over $\mathcal{H}$ is a normal algebraic subgroup of $\mathfrak{G}$, therefore $K = f(\mathfrak{G}(\mathcal{H}))$ is a normal algebraic subgroup of G. Now it is known (Chevalley and Kolchin [4], § 1, Th. 3) that there exists a matric representation g of G with kernel K; furthermore, although it is not specifically stated in [4], the proof there (§ 4) provides such a g which is a rational homomorphism of G. If we denote the image of g by H then (Chapter I, corollary to Proposition 5) H is an algebraic matric group. Clearly $g \circ f$ is a rational homomorphism of $\mathfrak{G}$ with kernel $\mathfrak{G}(\mathcal{H})$ and image H. Now (Chapter I, § 3, Example 1) the mapping r of $\mathfrak{G}$, which to each $\sigma \in \mathfrak{G}$ corresponds the restriction of σ

[3] Or even ([7], Ch. III, § 3, Th. 4, p. 797) a weakly normal extension.

to $\mathfrak{A}$, is a rational homomorphism of $\mathfrak{G}$ with kernel $\mathfrak{G}(\mathfrak{A})$ and image the Galois group $\mathfrak{H}$ of $\mathfrak{A}$ over $\mathfrak{F}$. Therefore (Chapter II, § 3, Proposition 1) there exists a birational isomorphism of $\mathfrak{H}$ onto H so that, by the present theorem, $\mathfrak{A}$ is a Picard-Vessiot extension of $\mathfrak{F}$.

THEOREM 3. *If the strongly normal extension $\mathfrak{G}$ of $\mathfrak{F}$ is of transcendence degree* 1 *and if $\mathfrak{F}$ is relatively algebraically closed in $\mathfrak{G}$ then there exists an element $\alpha \in \mathfrak{G}$ such that either α is primitive over $\mathfrak{F}$ and $\mathfrak{G} = \mathfrak{F}\langle\alpha\rangle$, or α is exponential over $\mathfrak{F}$ and $\mathfrak{G} = \mathfrak{F}\langle\alpha\rangle$, or α is weierstrassian over $\mathfrak{F}$ and $\mathfrak{G}$ is an abelian algebraic extension of $\mathfrak{F}\langle\alpha\rangle$ of finite degree. In the last case, if $\mathfrak{F}$ is algebraically closed then the weierstrassian element α may be chosen so that $\mathfrak{G} = \mathfrak{F}\langle\alpha\rangle$.*

This theorem is not new, being the main substance of Th. 7 of [7], Ch. III, p. 809. We include it here to show how it can be derived as a rather easy consequence of Theorem 1, and Theorem 1 of Chapter II; this makes it possible to avoid the long, tedious, and intricate computations of the proof in [7], pp. 809-823.

Proof. It follows from the hypothesis that the Galois group $\mathfrak{G}$ has dimension 1 and is irreducible; by Theorem 1 of Chapter II there exists a birational isomorphism c of $\mathfrak{G}$ onto a group variety G defined over $\mathfrak{C}$, which must also be of dimension 1 and irreducible. It is known (see Weil [14], § V, No. 38, pp. 68-69) that every such group variety is birationally isomorphic to (and therefore we may assume equal to) either the additive group of the field $\mathfrak{C}^*$ of all constants or the multiplicative group of $\mathfrak{C}^*$, or else a group variety which is a complete curve of genus 1 in projective space. In the last case we may always assume (for example, see Chevalley [2], Ch. II, §3) that the curve has an equation of the form $X_0X_2^2 = P(X_0, X_1)$, where P is a homogenous polynomial in $\mathfrak{C}[X_0, X_1]$ of degree 3 without multiple factor; applying a suitable transformation $(X_0, X_1, X_2) \to (X_0, aX_0 + bX_1, cX_2)$ if necessary, we lose no generality in supposing that

$$P(X_0, X_1) = 4X_1^3 - g_2X_0^2X_1 - g_3X_0^3,$$

where $g_2, g_3 \in \mathfrak{C}$. We may take the law of composition so that the group unity is $(0:0:1)$, which is the sole point at infinity on the curve, and such that the product $(1:\eta:\zeta)$ of two points $(1:\alpha:\beta)$ and $(1:a:b)$ with $\alpha \neq a$ is given by

$$\eta = -\alpha - a + 4^{-1}(\alpha - a)^{-2}(\beta - b)^2, \tag{3}$$

$$\zeta = -2^{-1}(\beta + b) + 2^{-1}3(\alpha - a)^{-1}(\beta - b) - 4^{-1}(\alpha - a)^{-3}(\beta - b)^3. \tag{4}$$

Consider the first case. By Theorem 1, there exists an element $\alpha \in \mathfrak{H}$ such that $c(\sigma) = \sigma\alpha - \alpha$ for every $\sigma \in \mathfrak{G}$. Since $c(\sigma) = 0$ only when σ is the identity, only the identity in $\mathfrak{G}$ leaves α invariant, so that $\mathfrak{H} = \mathfrak{F}\langle\alpha\rangle$. Applying the derivation operator δ_i to both sides of the equation $\sigma\alpha = \alpha + c(\sigma)$ we obtain $\sigma\delta_i\alpha = \delta_i\alpha (\sigma \in \mathfrak{G})$, so that $\delta_i\alpha \in \mathfrak{F}$ $(1 \leqq i \leqq m)$, and α is primitive over $\mathfrak{F}$.

Next consider the second case. By Theorem 1, there exists a nonzero element $\alpha \in \mathfrak{H}$ such that $c(\sigma) = \alpha^{-1}\sigma\alpha$ for every $\sigma \in \mathfrak{G}$. Since $c(\sigma) = 1$ only when σ is the identity, only the identity in $\mathfrak{G}$ leaves α invariant, so that $\mathfrak{H} = \mathfrak{F}\langle\alpha\rangle$. From the equation $\sigma\alpha = \alpha c(\sigma)$ we find that $\sigma\delta_i\alpha = (\delta_i\alpha)c(\sigma)$, whence $\sigma(\alpha^{-1}\delta_i\alpha) = \alpha^{-1}\delta_i\alpha$ $(\sigma \in \mathfrak{G})$, so that $\alpha^{-1}\delta_i\alpha \in \mathfrak{F}$ $(1 \leqq i \leqq m)$, and α is exponential over $\mathfrak{F}$.

Consider finally the third case. By Theorem 1, there exists a point $(1 : \alpha : \beta)$ on the curve $G^\dagger$, with $\alpha, \beta \in \mathfrak{H}$, and there exists an integer $n > 0$ $(n = 1$ if $\mathfrak{F}$ is algebraically closed), such that $c(\sigma)^n = (1 : \alpha : \beta)^{-1}$ $(1 : \sigma\alpha : \sigma\beta)$ for every $\sigma \in \mathfrak{G}$. Not every element of $\mathfrak{G}$ has order n, so that $c(\sigma)^n$ fails to be the identity for some $\sigma \in \mathfrak{G}$; therefore $(1 : \alpha : \beta)$ is not an invariant of the whole group $\mathfrak{G}$, hence is not rational over $\mathfrak{F}$. Since $\mathfrak{F}$ is relatively algebraically closed in $\mathfrak{H}$ and $\mathfrak{H}$ is of transcendence degree 1 over $\mathfrak{F}$, $\mathfrak{H}$ is an algebraic extension of $\mathfrak{F}\langle\alpha, \beta\rangle$ which clearly must be of finite degree, normal, and indeed abelian; if $\mathfrak{F}$ is algebraically closed, so that $n = 1$, then $\mathfrak{H} = \mathfrak{F}\langle\alpha, \beta\rangle$. For any $\sigma \in \mathfrak{G}$ which is not of order n (in particular, for a generic element of $\mathfrak{G}$) we may write $c(\sigma)^n = (1 : a : b)$, where a, b are constants. Then $\sigma\alpha$ and $\sigma\beta$ are given by the second members of (3) and (4), respectively. Making use of the fact that

$$\beta^2 = 4\alpha^3 - g_2\alpha - g_3, \tag{5}$$

these may be rewritten to obtain

$$\sigma\alpha = 4^{-1}(\alpha - a)^{-2}[4a\alpha^2 + (4a^2 - g_2)\alpha - (g_2a + 2g_3) - 2b\beta], \tag{6}$$

$$\begin{aligned}\sigma\beta = 4^{-1}(\alpha - a)^{-3}[&(4\alpha^3 + 12a\alpha^2 - 3g_2\alpha - g_2a - 4g_3)b \\ &+ ((-12a^2 + g_2)\alpha - 4a^3 + 3g_2a + 4g_3)\beta].\end{aligned} \tag{7}$$

From (5) we find $2\beta\delta_i\beta = (12\alpha^2 - g_2)\delta_i\alpha$. Therefore application of δ_i to (6) yields

$$\begin{aligned}\sigma\delta_i\alpha &= 4^{-1}(\alpha - a)^{-2}((8a\alpha + 4a^2 - g_2)\delta_i\alpha - 2b)\delta_i\beta \\ &\quad - 2^{-1}(\alpha - a)^{-3}(4a\alpha^2 + (4a^2 - g_2)\alpha - g_2a - 2g_3 - 2b\beta)\delta_i\alpha \\ &= 4^{-1}(\alpha - a)^{-2}((8a\alpha + 4a^2 - g_2)\beta - b(12\alpha^2 - g_2))\beta^{-1}\delta_i\alpha \\ &\quad - 2^{-1}(\alpha - a)^{-3}((4a\alpha^2 + (4a^2 - g_2)\alpha - g_2a - 2g_3)\beta \\ &\qquad\qquad - 2b(4\alpha^3 - g_2\alpha - g_3))\beta^{-1}\delta_i\alpha.\end{aligned}$$

The last member equals $\beta^{-1}\delta_i\alpha$ multiplied by the second member of (7). Therefore $\sigma\delta_i\alpha = (\sigma\beta)\beta^{-1}\delta_i\alpha$, that is $\sigma(\beta^{-1}\delta_i\alpha) = \beta^{-1}\delta_i\alpha$ $(\sigma \in \mathfrak{G})$, so that $\beta^{-1}\delta_i\alpha \in \mathfrak{F}$ $(1 \leqq i \leqq m)$. It follows that $\mathfrak{F}\langle\alpha, \beta\rangle = \mathfrak{F}\langle\alpha\rangle$, and also (with the help of (5) that $(4\alpha^3 - g_2\alpha - g_3)^{-1}$ $(\delta_i\alpha)^2$ is the square of an element of $\mathfrak{F}$ $(1 \leqq i \leqq m)$, that is, that α is weierstrassian over $\mathfrak{F}$.

COLUMBIA UNIVERSITY
AND UNIVERSITY OF PARIS.

REFERENCES.

[1] N. Bourbaki, *"Algèbre,"* Chapters IV-V (Actualités scientifiques et industrielles 1102), Hermann et Cie., Paris, 1950.

[2] C. Chevalley, *"Introduction to the theory of algebraic functions of one variable"* (Mathematical Surveys VI), American Mathematical Society, New York, 1951.

[3] ———, *"Théorie des groupes de Lie, Tome II: Groupes algébriques"* (Actualités scientifiques et industrielles 1152), Hermann et Cie., Paris, 1951.

[4] ———, and E. Kolchin, "Two proofs of a theorem on algebraic groups," *Proceedings of the American Mathematical Society*, vol. 2 (1951), pp. 126-134.

[5] E. R. Kolchin, "Algebraic matric groups and the Picard-Vessiot theory of homogeneous linear ordinary differential equations," *Annals of Mathematics*, vol. 49 (1948), pp. 1-42.

[6] ———, "Picard-Vessiot theory of partial differential fields," *Proceedings of the American Mathematical Society*, vol. 3 (1952), pp. 596-603.

[7] ———, "Galois theory of differential fields," *American Journal of Mathematics*, vol. 75 (1953), pp. 753-824.

[8] H. Matsumura, "Automorphism-groups of differential fields and group varieties," *Memoirs of the College of Science, University of Kyoto, Series A*, vol. 28 (1954), pp. 283-292.

[9] S. Nakano, "On invariant differential forms on a group variety," *Journal of the Mathematical Society of Japan*, vol. 2 (1951), pp. 216-227.

[10] ———, "Note on group varieties," *Memoirs of the College of Science, University of Kyoto, Series A*, vol. 27 (1952), pp. 55-66.

[11] M. Rosenlicht, "Generalized Jacobian varieties," *Annals of Mathematics*, vol. 59 (1954), pp. 505-530.

[12] J.-P. Serre, "Faisceaux algébriques cohérents," *Annals of Mathematics*, vol. 61 (1955), pp. 197-278.

[13] A. Weil, *"Foundations of Algebraic geometry"* (American Mathematical Society Colloquium Publications, vol. 29), New York, 1946.

[14] ———, "Variétés abéliennes et courbes algébrique (Actualités scientifiques et industrielles 1064), Hermann et Cie., Paris, 1948.

ENGEL RINGS AND A RESULT OF HERSTEIN AND KAPLANSKY.*

By M. P. Drazin.

---

1. Introduction. It is known that if, in an associative ring R, some power of every element is central, then the commutator ideal R_1 of R (i.e. the two-sided ideal of R generated by all the commutators $[a, b] = ab - ba$ with a, b in R) is nil; more precisely, if to each element y of R corresponds a positive integer $n = n(y)$ such that $[x, y^n] = 0$ for all x in R, then to each element c of R_1 corresponds a positive integer j such that $c^j = 0$. Special cases of this were established by Jacobson, Hua and others, after which Kaplansky [15] showed in the general case that R_1 at any rate lies in the Jacobson radical; that R_1 is actually nil was finally proved by Herstein [9]. It will be convenient to call a given ring R *commutator-nil* whenever its commutator ideal R_1 is nil.

Consider now another class of rings. For any x, y in R, define a sequence of elements $e_k(x, y)$ of R inductively by the rules

$$e_0(x, y) = x, \qquad e_k(x, y) = [e_{k-1}(x, y), y] \qquad (k = 1, 2, \cdots);$$

equivalently, in binomial notation,

$$e_k(x, y) = \sum_{j=0}^{k} (-1)^j C_j^k y^j x y^{k-j} \qquad (k = 0, 1, \cdots).$$

If to each pair x, y there corresponds an integer $k = k(x, y)$ such that $e_k(x, y) = 0$, then we call R an *Engel ring*. The present paper grew from an investigation of the properties of the commutator ideals of arbitrary Engel rings. The relevance of the above-mentioned result of Herstein and Kaplansky soon appeared; indeed, it became clear that the two kinds of rings can profitably be regarded as special cases of another type, the *K-rings*, where we say that a given ring R is a K-ring if, to each pair of elements x, y of R, there correspond positive integers $k = k(x, y)$ and $n = n(x, y)$ such that $e_k(x, y^n) = 0$. The rings studied by Herstein and Kaplansky are obtained on imposing the restrictions that $k = 1$ for all x, y and that n be independent of x, while of course the Engel rings are just those K-rings with $n = 1$ for all x, y.

---

* Received June 9, 1955.

Two of our principal results (both generalizing Herstein's) are that a given K-ring must always be commutator-nil whenever its index functions k, n are both functions of y alone, and whenever $k = 1$ for all x, y. Since every nil ring is a K-ring (and indeed falls in all of the special classes of K-rings we have mentioned), to prove without further hypothesis that R_1 is more than merely nil would imply a non-trivial statement about the commutator-ideal of an arbitrary nil ring. We establish no such statement here, being instead more concerned in our later results with extending the commutator-nil property (or even slightly weaker properties) to wider classes of rings. For example, we prove that the commutator ideal of an arbitrary K-ring is always contained in the Jacobson radical.

For the proofs of the facts, we need two auxiliary results about the structure of (the semi-simple part of) an arbitrary K-ring. These are derived in a more general context in Section 2. We define, for each positive integer m, a class of rings called m-rings. These classes form an expanding sequence, i.e., if l, m are any given integers satisfying $1 \leqq l \leqq m$, then every l-ring is an m-ring. Every K-ring is a 1-ring (and hence an m-ring for all m); also every algebraic ring of bounded index, and every ring with minimal condition on its left (or right) ideals is an m-ring for some m.

Four successively narrower classes of rings have now been mentioned: the m-rings (m arbitrary), the 1-rings, the K-rings and the Engel rings. Having discussed the first and widest class in Section 2 we proceed step by step from the general to the particular, examining the other three classes in turn in Sections 3, 4 and 5. In Section 3 we consider the 1-rings (which include, besides the K-rings, all strongly regular rings and another class of rings investigated recently by Herstein [10] and others), restating the specialization $m = 1$ of the results of Section 2, and proving also one other result. Our theorems on K-rings as such are proved in Section 4. In Section 5 we press our investigation a little further for the Engel rings, showing that certain of these satisfy conditions stronger than the commutator-nil property.

In our final section we consider those rings in which, for each pair of elements x, y, there are integers k, m, n such that $e_k(x^m, y^n) = 0$. These include all the K-rings as the special case $m = 1$, and we show that many of our results for K-rings can be extended to this more general class.

Throughout, R will denote an associative ring, not necessarily with a unit element, admitting a commutative ring F of operators (i.e. endomorphisms α of the additive group of R, subject to $\alpha(xy) = (\alpha x)y = x(\alpha y)$ for all $x, y \in R$); we may suppose without loss of generality that F contains the

identity operator. The case of a "ring without operators" is included in this scheme on taking F to be just the ring of integers (or an appropriate quotient ring). When we refer to subrings (etc.) of R these should always be understood as sub-F-rings (etc.), i.e. as being mapped into themselves by every operator in F.

2. Two theorems about m-rings. Given any ring R over F and any positive integer m, we shall say that R is an *m-ring* whenever the following statement is true: to each choice of x, y in R there correspond $\alpha = \alpha(x,y)$ in F, $a = a(x,y)$ in R and integers $k = k(x,y)$, $l = l(x,y)$, $n = n(x,y)$ such that

$$1 \leqq l \leqq m, \qquad x^{m-l} e_k(\alpha x^{l+1} + x^{l+1}a - x^l, y^n) = 0.$$

Every algebraic ring of bounded index, and every ring with minimal condition on its left or right ideals, is of this type for some m (even with $k = 0$, $l = m$, $n = 1$). We establish now two results about the structure of the semi-simple parts of arbitrary m-rings, interpreting semi-simplicity in the sense of Koethe for Theorem 2.1, and in the sense of Jacobson for Theorem 2.2.

THEOREM 2.1. *Let R be a given m-ring without non-zero nil two-sided ideals. Then every nilpotent element z of R satisfies $z^{m^2} = 0$.*

Proof. The main step of the proof is to show that if a given element x of R satisfies $x^{m+1} = 0$, then x^m generates a nil two-sided ideal (and must consequently vanish).

We show first that, for any $b \in R$, bx^m is nilpotent. We have, for any $\alpha \in F$, any $a \in R$ and any positive integers k, l, n with $l \leqq m$,

$$x^{m-l} e_k(\alpha x^{l+1} + x^{l+1}a - x^l, (bx^m)^n)$$
$$= x^{m-l}(\alpha x^{l+1} + x^{l+1}a - x^l)(bx^m)^{nk} = -x^m(bx^m)^{nk};$$

but R is an m-ring, so this, and consequently $(bx^m)^{nk+1}$, vanishes for some choice of k, n.

We can now deduce that x^m generates a nil two-sided ideal. The typical element of this ideal has the form $d = \beta x^m + bx^m + x^m c$, where $\beta \in F$ and $b, c \in R$. Using the fact that $bx^m \cdot x^m = x^m \cdot x^m c = 0$, we see that, for each positive integer j,

$$d^j = \beta \sum_{i=0}^{j-1} (x^m c)^i x^m (bx^m)^{j-i-1} + \sum_{i=0}^{j} (x^m c)^i (bx^m)^{j-i}$$
$$= \beta x^m \sum_{i=0}^{j-1} (cx^m)^i (bx^m)^{j-i-1} + (bx^m)^j + x^m \sum_{i=1}^{j} (cx^m)^{i-1} c (bx^m)^{j-i}.$$

Also, by what we have already proved, there are positive integers p, q such that $(bx^m)^p = (cx^m)^q = 0$, whence it follows that $d^{p+q} = 0$.

We have now shown that $x^{m+1} = 0$ implies $x^m = 0$. Let z be an arbitrary nilpotent element of R, and let r be the smallest positive integer for which $z^{r^m} = 0$. Then, if $r > m$, we should have $(z^{r-1})^{m+1} = z^{r-m-1}z^{rm} = 0$, whence, on taking $x = z^{r-1}$, it would follow that $(z^{r-1})^m = 0$, contrary to our choice of r. Consequently $r \leqq m$, i.e. $z^{m^2} = 0$ as required.

We cannot, even for the 1-rings, hope to replace the hypothesis of no nil ideals in Theorem 2.1 by the weaker one of no nilpotent ideals; for Baer [3; 540-543] has constructed a non-zero nil ring having no non-zero nilpotent ideals. Indeed, by Nagata [18] there exist nil rings (and *a fortiori* 1-rings and m-rings), having no non-zero nilpotent ideals, whose nilpotent elements are of unbounded index.

We have stated Theorem 2.1 in a way which avoids explicit reference to the Koethe radical (i.e. the union of all two-sided nil ideals of R). An equivalent statement is the following: If R is any given m-ring, with Koethe radical N, say, then, given any $z \in R$, the existence of an integer s with $z^s \in N$ implies $z^{m^2} \in N$ (in particular, R/N has no non-zero nilpotent elements if $m = 1$). If J denotes the Jacobson radical of R, then an analogous statement holds with N replaced by J (since R/J has no non-zero nil ideals).

There being regrettably few problems which yield to attacks based on the rather weak structure theory associated with the Koethe radical, we next prove another auxiliary theorem for use when the Jacobson theory is appropriate. This generalizes a result of Arens and Kaplansky [2; Theorem 3.1]; to describe their result in our terminology, they considered only m-rings in which $k = 0$ for all x, y.

THEOREM 2.2. *Let R be any given primitive m-ring over F. Then R is isomorphic to the complete ring $M_q(D)$ of $q \times q$ matrices over D, where $q \leqq m$ and D is a suitable division ring over F.*

Proof. R, being primitive, is isomorphic to a dense ring M of linear transformations of a vector space V over some division ring D; we shall denote the result of operating on $v \in V$ with $x \in M$ by vx, i.e. regard M as operating on V from the right.[1] We have only to show that V cannot contain $m+1$ elements $v_0, v_1, \cdots, v_m$ independent with respect to D.

In the contrary case, since M is dense, we could choose x, y in M so that

[1] This is dictated by our having used the expression $\alpha x^{i+1} + x^{i+1}a - x^i$ rather than $\alpha x^{i+1} + ax^{i+1} - x^i$ when defining m-rings (so as to ensure that every strongly regular ring is self-evidently a 1-ring).

$$v_h x = v_{h+1} \qquad (h = 0, 1, \cdots, m-1), \qquad v_m x = 0,$$

$$v_h y = 0 \qquad (h = 0, 1, \cdots, m-1), \qquad v_m y = v_m.$$

Then, for any $\alpha \in F$, any $a \in R$ and any $k \geqq 0$, $1 \leqq l \leqq m$, $n \geqq 1$,

$$v_0 x^{m-l} e_k(\alpha x^{l+1} + x^{l+1}a - x^l, y^n) = v_{m-l} e_k(\alpha x^{l+1} + x^{l+1}a - x^l, y^n)$$

$$= v_{m-l}(\alpha x^{l+1} + x^{l+1}a - x^l)(y^n)^k$$

$$= v_{m-l}x^{l+1}(\alpha y^{nk} + a y^{nk}) - v_{m-l}x^l y^{nk} = 0 - v_m y^{nk} = -v_m \neq 0$$

(by the D-independence of $v_0, v_1, \cdots, v_m$), contrary to the m-ring hypothesis.

Remark. The briefest examination of the proof of Theorem 2.1 shows that the full force of the m-ring hypothesis is not needed, but only its consequence that, for any b, x in R, $x^{m+1} = 0$ implies bx^m nilpotent. This will always be the case whenever, for each choice of b, x in R, a positive integer s exists such that $(bx^m)^s$ is in the two-sided ideal generated by x^{m+1}; and this is in turn entailed (on taking $y = bx^m$) by the hypothesis that, for each choice of x, y in R, an integer $t = t(x, y)$ exists such that $x^m y^t$ is in the two-sided ideal generated by yx and x^{m+1} (which condition is still satisfied by every m-ring, with $t(x, y) = k(x, y)n(x, y)$). These trivial comments are relevant to certain generalizations of Theorem 2.2 which will be described in detail elsewhere [6].

3. Remarks on the case $m = 1$. To say that R is a 1-ring means that, to each choice of x, y in R, there correspond α in F, a in R and integers k, n such that $e_k(\alpha x^2 + x^2 a - x, y^n) = 0$. It is not hard to see that, by replacing k and n respectively by

$$\max\{k(x, y), k(\alpha(x, y)x^2 + x^2 a(x, y) - x, y)\}$$

and

$$n(x, y)n(\alpha(x, y)x^2 + x^2 a(x, y) - x, y),$$

and by modifying a appropriately, we may always take $\alpha = 0$ in a 1-ring.

Applying Theorem 2.1 to R/N with $m = 1$, we have at once

THEOREM 3.1. *In any 1-ring, every nilpotent element generates a nil two-sided ideal.*

Analogously, it is easy to see that, in any commutator-nil ring, every nilpotent element generates a nil two-sided ideal. However, not every 1-ring is commutator-nil (consider for example any non-commutative division ring); and one can also verify without difficulty (e.g. from Theorem 3.3 below) that not every commutator-nil ring is a 1-ring.

Next, for any ring R, with Jacobson radical J, R/J is semi-simple in Jacobson's sense, so that R/J is isomorphic to a subdirect sum of primitive rings, each of which is a homomorphic image of R/J and hence of R. Consequently, by Theorem 2.2 and the homomorphism-invariance of the 1-ring property, we have

THEOREM 3.2. *If R is any 1-ring, with Jacobson radical J, then R/J is a subdirect sum of division rings.*

This may be regarded as a generalization of a familiar property of strongly regular rings (which have $J=0$); it also generalizes an argument of Herstein [10; Theorem 5].

Our final result in this section is included chiefly for the sake of its connexion with the existence of annihilators of ideals in rings with minimal condition on left (or right) ideals (cf. [5]). Of course every K-ring is a 1-ring (with $\alpha=0$, $a=0$), and clearly every idempotent element of a K-ring is central; for if $y^2=y$ then, for any x, n,

$$e_3(x,y^n)=e_3(x,y)=xy^3-3yxy^2+3y^2xy-y^3x=xy-yx,$$

so that $xy-yx=e_{2j+1}(x,y^n)$ $(j,n=1,2,\cdots)$. We show (cf. Herstein [10; Lemma 4]) that this result in fact holds for all 1-rings:

THEOREM 3.3. *In any 1-ring, every idempotent element is central.*

Proof. Let R be a given 1-ring and y any idempotent in R; we must show that, for any given $b\in R$, $by=yb$.

Define $x=by-yby$, so that $yx=0$ and $xy=x$, whence $x^2=0$ and consequently (since R is a 1-ring) integers k, n exist satisfying

$$0=e_k(-x,y^n)=-e_k(x,y)=-\sum_{j=0}^{k}(-1)^j C_j^k y^j x y^{k-j}=-xy^k=-x,$$

i.e. $yby=by$; and a similar argument shows that $yby=yb$, as required.

The above proof requires only that $y^2=y$, together either with $yx=0$ and $xy=x$, or with $xy=0$ and $yx=x$, implies $x=0$. This holds in any ring R admitting a positive integral-valued function $t(x,y)$ such that, for all $x,y\in R$, (a) xy^t is in the two-sided ideal generated by yx and x^2, and (b) y^tx is in the two-sided ideal generated by xy and x^2; here (a) is just the specialization $m=1$ of the condition mentioned at the end of Section 2, while (b) (obtainable from (a) by writing all products in reverse order) is also satisfied by every 1-ring.

It is easy to deduce from Theorems 3.1 and 3.3 (by observing that every element of R/N is a linear combination of "principal idempotent

elements") that every 1-ring which is also an algebraic algebra over an algebraically closed field must be commutator-nil; however, for the special case of K-rings, this result is included in Theorem 4.5 below.

4. The K-ring theorems. The proof of our first main result is elementary, and self-contained apart from appeals to Theorem 3.1 and to the result of Herstein mentioned in the Introduction.

LEMMA 4.1. *Let R be a ring in which, to each element y, there correspond positive integers $k = k(y)$, $n = n(y)$ such that $e_k(x, y^n) = 0$ for all x in R; suppose also that R has characteristic either* (i) *zero, or* (ii) *a prime. Then R is commutator-nil.*

Proof. (i) We need only consider rings without non-zero nilpotent elements. For suppose the result were known for these, and let N denote the union of all two-sided nil ideals of R. Then (as is well known and easy to see) N is itself a two-sided nil ideal of R, and R/N has no non-zero two-sided nil ideals. Also R/N inherits the Engel hypothesis on R and, by Theorem 3.1, is consequently even a ring without non-zero nilpotent elements; we observe further that (since N is nil) R/N has zero characteristic. Hence, if the result were known for rings of zero characteristic without nilpotent elements, then R/N would be commutative, i.e. R_1 would be nil.

Thus we may assume without loss of generality that R itself has no non-zero nilpotent elements. Given $y, n(y)$, we may suppose also that $h = k(y)$ has been chosen as small as possible subject to $h \geqq 1$; we shall show that $h = 1$. In analogy with Leibniz's formula in differential calculus we have, for any non-negative integer j and any $a, b, c \in R$,

$$e_j(ab, c) = \sum_{i=0}^{j} C_i^j e_i(a, c) e_{j-i}(b, c);$$

in particular, for any $x \in R$,

$$e_{2h-2}(x^2, y^n) = \sum_{i=0}^{2h-2} C_i^{2h-2} e_i(x, y^n) e_{2h-2-i}(x, y^n) = C_{h-1}^{2h-2} e_{h-1}^2(x, y^n).$$

Now, if $h > 1$, then $2h - 2 \geqq h$ and so the expression on the left here vanishes, whence (since R is of zero characteristic and contains no non-zero nilpotent elements) it would follow that $e_{h-1}(x, y^n) = 0$ for all x in R, contrary to our choice of h. Hence in fact $h = 1$, i.e. we have $[x, y^{n(y)}] = 0$ for all x, y in R; the result now follows from Herstein [9].

(ii) Suppose that, for some prime p, $pR = 0$. Since the binomial coefficients $C_j^{p^r}$ $(j = 1, \cdots, p^r - 1; r = 1, 2, \cdots)$ are all divisible by p, we have, for all $x, z \in R$ and every positive integer r,

$$e_{p^r}(x,z) = \sum_{j=0}^{p^r} (-1)^j C_j^{p^r} z^j x z^{p^r-j} = xz^{p^r} + (-1)^{p^r} z^{p^r} x = [x, z^{p^r}]$$

(on considering separately the cases $p=2$, $p \neq 2$). Also, by our hypothesis on R, $e_{p^r}(x, y^n) = 0$ whenever $p^r \geqq k(y)$; thus, taking $z = y^n$, we see that, for all sufficiently large r (the criterion of sufficiency being independent of x), $[x, y^{r(y)p^r}] = 0$ for all $x \in R$. The result now follows from Herstein's as before.

We have next to remove the restrictions (i) and (ii) from Lemma 4.1. This is of course immediate when F is a field, and will now be done in the general case by an argument communicated to the writer by A. S. Amitsur; he naturally proved his result in the special form immediately applicable to our present purpose, but his argument seems interesting enough to deserve formal statement in somewhat more general terms. Calling an arbitrary subset Q of a ring nil whenever every element of Q is nilpotent, we have

LEMMA 4.2. *Let $\mathfrak{R}$ be any given set of rings, in each member R of which is given also a subset $Q(R)$; suppose further that, given any $R \in \mathfrak{R}$ and any homomorphism θ of R, $\mathfrak{R}$ is such that $R\theta \in \mathfrak{R}$ and $Q(R)\theta \leqq Q(R\theta)$. Then, if also $Q(R)$ is nil for all those $R \in \mathfrak{R}$ having characteristic either* (i) *zero or* (ii) *a prime, $Q(R)$ must in fact be nil for every $R \in \mathfrak{R}$.*

Proof. Let R be an arbitrary member of $\mathfrak{R}$, and c any element of $Q(R)$; we have to prove c nilpotent.

Let A denote the subset of R consisting of all elements of finite non-zero additive order (i.e. the "torsion subgroup" of the additive group of R). Clearly A is a two-sided ideal of R, and so R/A exists; also R/A has zero characteristic and is of course a homomorphic image of R. Hence, by (i), $Q(R/A)$ is nil, and so, by our hypothesis $Q(R)\theta \leqq Q(R\theta)$, some power c^i of c lies in A; in other words we can find positive integers i, r so that $rc^i = 0$. Let $s = s(c)$ denote the smallest positive integer annihilating a power of c, say $sc^j = 0$; we have then only to show that $s = 1$.

In the contrary case, we could find a prime p dividing s, say $s = pt$ where $1 \leqq t < s$. Since pR is a two-sided ideal of R, (ii) tells us that $Q(R/(pR))$ is nil; thus, on using once more our hypothesis $Q(R)\theta \leqq Q(R\theta)$, we see that some power of c lies in pR, say $c^k = pd$, where $d \in R$. But then

$$tc^{j+k} = ptc^j d = sc^j d = 0,$$

where $t < s$ is a positive integer, contrary to our choice of s; the result follows from this.

Now, by Lemma 4.1, on taking $Q(R) = R_1$ all the conditions of Lemma 4.2 are satisfied by the set $\mathfrak{R}$ of all rings over each of which are defined functions $k = k(y)$, $n = n(y)$ such that $e_k(x, y^n) = 0$ for all x, y. Thus we have proved

THEOREM 4.1. *If R is any ring in which, to each element y, there correspond positive integers $k = k(y)$, $n = n(y)$ such that $e_k(x, y^n) = 0$ for all x, then R is commutator-nil.*

It will appear in Theorem 6.2 below that a similar result holds when it is given that $e_{k(x)}(x, y^{n(x)}) = 0$ for all x, y.

The proving of Theorem 4.1 has been something of a cooperative piece of work. As already stated, the argument of Lemma 4.2 is due to Amitsur, while the proof of Lemma 4.1 owes much to some (unpublished) work of P. M. Cohn and K. W. Gruenberg; the writer is glad to acknowledge his debt to these three. Cohn and Gruenberg considered only Engel rings R with $k(x, y)$ independent of x; Gruenberg had proved Theorem 3.2 for these rings, and this enabled them to prove $R_1 \leqq J$ by reduction to the division ring case. Gruenberg himself settled the subcase of prime characteristic by an argument, suggested by [12; p. 102], of which our proof of Lemma 4.1 (ii) is only a trivial extension; and Cohn, without having Theorem 3.1 to argue from, dealt with the subcase of zero characteristic by an argument rather similar to that used in the second part of our proof of Lemma 4.1 (i). For division rings these two subcases exhaust the possibilities, so that no bridging argument along the lines of Lemma 4.2 was called for. Our Theorem 4.1 improves on Cohn and Gruenberg's result in two directions, and its proof appeals to Jacobson's structure theory only indirectly (through Herstein); we extend their result in a different way in Theorem 4.4 below.

Two published results, both included in Theorem 4.1, should also be mentioned. The case of Engel rings with k independent of both x and y was settled by Amitsur [1]; indeed; he showed that in such a ring R the left (or right) ideal generated by any given element of R_1 is nil of bounded index (this was stated in his paper only for algebras over infinite fields, but he has pointed out in a letter that this restriction is inessential). His work in turn superseded a result of Jennings [14; Theorem 1], who showed that R_1 is nil whenever a fixed positive integer h exists such that the repeated commutator $[[\cdots[[x_1, x_2], x_3], \cdots], x_h] = 0$ for all choices of $x_1, \cdots, x_h$ in R.

Theorem 4.1 does not fully answer the question whether every K-ring is commutator-nil; and the writer knows no counter-example to this proposition. Before discussing how the proof of Theorem 4.1 might be extended to deal with arbitrary K-rings (i.e. to allow k, n to depend also on x), we first note a slight generalization of Herstein's result:[2]

[2] *Added in proof.* Shortly after the present paper was submitted Herstein himself published a proof of the same generalization in the *Canadian Journal of Mathematics*, vol. 7 (1955), pp. 411-412.

THEOREM 4.2. *If in a given ring R there corresponds to each pair of elements x, y a positive integer $n = n(x, y)$ such that $xy^n = y^n x$, then R is commutator-nil.*

Proof. We observe that, since $\alpha[a, b] = [\alpha a, b]$, $\alpha d[a, b] = d[\alpha a, b]$ for any admissible operator α and any a, b, d in R, the ideal R_1 does not depend on the given set of operators on R; thus we may regard R as a ring without operators. We reduce the problem to Herstein's by showing that R is locally of the type he considered.

Let c be any element of R_1, i.e. any sum of terms of the forms $[a_i, b_i]$, $d_i[a_i, b_i]$, $[a_i, b_i]f_i$, $d_i[a_i, b_i]f_i$. Let the a_i, b_i, d_i, f_i involved in such an expression for c be enumerated as $x_1, \cdots, x_t$ (order being irrelevant), and let X denote the ring they generate. Then c is in the commutator ideal of X, and so it will be enough to prove that X is commutator-nil.

Take any $y \in R$. By the hypothesis on R, $[x_i, y^{n(x_i, y)}] = 0$ $(i = 1, \cdots, t)$, and so, if we define $n(y) = n(x_1, y) \cdots n(x_t, y)$, then $[x_i, y^{n(y)}] = 0$ $(i = 1, \cdots, t)$. Since the x_i generate X, it follows that $[x, y^{n(y)}] = 0$ for all $x \in X$, $y \in R$. In particular, for each $y \in X$, $y^{n(y)}$ lies in the centre of X; hence, by Herstein's result [9], X is commutator-nil as required.

Unfortunately it does not seem possible, even when we are given that k is independent of both x and y, to remove the restriction $n = n(y)$ in Theorem 4.1 by the above sort of argument. However, using Theorem 4.2 and the argument of Lemma 4.1 (ii), we can at once deduce

THEOREM 4.3. *Every K-ring of prime characteristic is commutator-nil.*

To prove that *all* K-rings are commutator-nil (i.e. to remove the restrictions $k = k(y)$, $n = n(y)$ from Theorem 4.1) it would, in virtue of Lemma 4.2 and Theorem 4.3, be enough to consider only rings of zero characteristic, i.e. to remove the restrictions $k = k(y)$, $n = n(y)$ from Lemma 4.1 (i). And we can of course still use Theorem 3.1 to reduce further to rings without non-zero nilpotent elements (for which one would have to prove commutativity). In other words, the K-rings of zero characteristic without non-zero nilpotent elements are critical for the whole question whether all (and, if not, which) K-rings are commutator-nil; it would be particularly interesting to know the answer for the subclass of Engel rings.

The results of this section have so far applied only to certain special kinds of K-ring. We now establish a property common to all K-rings:

THEOREM 4.4. *Every K-ring is commutative modulo its Jacobson radical.*

Proof. By Theorem 3.2, we may confine ourselves to showing that, if a

given division ring D is also a K-ring, then it is commutative; and, by Theorem 4.3, we may suppose that D has zero characteristic. By Theorem 4.2, it will be enough to prove that, if $k(x,y)$, $n(x,y)$ are the given index functions of D as a K-ring, then $[x, y^{n(x,y)}] = 0$ for all $x, y \in D$. Suppose by way of contradiction that D contained x, y such that $[x, y^{n(x,y)}] \neq 0$. On writing $y^{n(x,y)} = z$ this becomes $[x,z] \neq 0$, while we know that $e_{k(x,y)}(x,z) = 0$, so there is a positive integer j $(\leqq k(x,y) - 1)$ such that $e_j(x,z) \neq 0$, $e_{j+1}(x,z) = 0$. Since j is positive, $b = e_{j-1}(x,z)$ exists, and, writing $c = e_j(x,z)$, we have $c = [b,z] \neq 0$, $[c,z] = 0$. Since $c \neq 0$, c^{-1} exists and commutes with z (since c does); hence, defining $d = -c^{-1}b$, we find $[z,d] = c^{-1}[z,-b] = 1$. Writing $f = zd$, we then have $[z,f] = z \neq 0$.

We shall obtain our contradiction by showing that the relation $[z,f] = z \neq 0$ cannot hold between elements z, f of a division K-ring D having zero characteristic. We note that

$$[z, f^q] = \{(f+1)^q - f^q\}z \qquad (q = 1, 2, \cdots);$$

for the case $q = 1$ of this merely states that $[z,f] = z$, while, if the relation holds for a given q, then

$$\begin{aligned}[z, f^{q+1}] &= f^q[z,f] + [z,f^q]f = f^q[z,f] + \{(f+1)^q - f^q\}zf \\ &= f^q[z,f] + \{(f+1)^q - f^q\}([z,f] + fz) \\ &= (f+1)^q z + \{(f+1)^q - f^q\}fz = \{(f+1)^{q+1} - f^{q+1}\}z.\end{aligned}$$

It then follows, by another trivial induction, that

$$e_i(z, f^q) = \{(f+1)^q - f^q\}^i z \qquad (i = 0, 1, \cdots; q = 1, 2, \cdots),$$

and indeed, since $[z, f+h] = [z,f] = z$, that

$$e_i(z, (f+h)^q) = \{(f+h+1)^q - (f+h)^q\}^i z \qquad (h, i = 0, 1, \cdots; q = 1, 2, \cdots).$$

Since D is a K-ring, we can find positive integers $k = k(z,f)$, $n = n(z,f)$ such that $e_k(z, f^n) = 0$, whence (since $z \neq 0$) we deduce $(f+1)^n = f^n$. But this implies

$$(f+h)^n - f^n = 0 \qquad (h = 1, 2, \cdots);$$

for we have just proved the case $h = 1$, while the result for given h entails

$$\begin{aligned}\{(f+h+1)^n - f^n\}^k z &= \{(f+h+1)^n - (f+h)^n\}^k z \\ &= e_k(z, (f+h)^n) = e_k(z, f^n) = 0.\end{aligned}$$

Now the n equations $(f+h)^n - f^n = 0$ $(h = 1, 2, \cdots, n)$, regarded as homo-

geneous linear relations between the n quantities $C_0{}^n, C_1{}^n f, \cdots, C_{n-1}{}^n f^{n-1}$, have (Vandermonde) determinant $\Delta = 1!2!\cdots n!$; we conclude that

$$\Delta C_0{}^n = \Delta C_1{}^n f = \cdots = \Delta C_{n-1}{}^n f^{n-1} = 0,$$

which (since D has characteristic zero) gives our sought-for contradiction. Thus in fact $[x, y^{n(x,y)}] = 0$ for all x, y, and the result follows.

In the Engel case the above proof reduces almost to triviality (and is in particular independent of Herstein's result and of characteristic considerations). From well-known properties of the Jacobson radical, we have immediately

THEOREM 4.5. *Every π-regular K-ring is commutator-nil. Every K-ring with minimal condition on its left (or right) ideals has nilpotent commutator ideal.*

5. Engel rings. The results of Section 4 can be extended either in the direction of proving something stronger than the commutator-nil property for suitably selected special K-rings, or by showing that the commutator-nil property (or commutativity modulo J) in fact holds under hypotheses less stringent than those made above. We shall obtain a few such extensions in this section and the next.

In view of the result of Jennings quoted shortly before Theorem 4.2 above, perhaps the most conspicuous problem under the first heading is to determine for which K-rings R an integer h exists such that $[[\cdots[[x_1, x_2], x_3], \cdots], x_h] = 0$ for all choices of $x_1, \cdots, x_h$ in R (conversely, of course the existence of such an h implies that R is a K-ring, and indeed an Engel ring with $k = h - 1$ independent of x and y). Let $[R]$ denote the Lie ring associated with R, i.e. having the same additive group as R, but with the multiplication $[x, y]$. Then our problem may be restated as follows: when is the K-ring condition on R equivalent to the nilpotence of $[R]$ (as a Lie ring)?

It is known that every finite-dimensional Engel algebra R has $[R]$ nilpotent. Also, even in the wider class of arbitrary Lie rings L (i.e. not only those which come from associative rings in the way just described), it is known that the Engel condition $e_k(x, y) = 0$, with index function $k = k(x, y)$, implies the nilpotence of L whenever either (1) L is soluble and finitely generated, or (2) L is soluble and of zero characteristic and $k(x, y)$ is independent of both x and y, or (3) L satisfies the maximal condition on (Lie) subrings and $k(x, y)$ is independent of x. These three results are respectively due to Gruenberg [8; Theorem 1'], Higgins [11; Theorem 1] and Zorn [19];

Higgins also showed [Theorem 4] that zero characteristic and the Engel condition with $k(x,y) \leqq 4$ together always ensure solubility. We now ask how far these provisos may be dispensed with when L is the Lie ring $[R]$ of some associative ring R.

We certainly cannot hope to prove that every associative Engel ring R has $[R]$ nilpotent: for Cohn [4] has constructed, for each prime p, an associative ring of characteristic p in which $e_{p+1}(x,y)$ vanishes identically but which contains an infinite sequence of elements x_j such that $[[\cdots[x_1,x_2],\cdots],x_j] \neq 0$ $(j=2,3,\cdots)$. However, we do at any rate have

THEOREM 5.1. *Let R be any given associative Engel ring, with index function $k(x,y)$ independent of both x and y, and suppose further that R has zero characteristic and satisfies the minimal or maximal condition on left (or right) ideals; then the Lie ring $[R]$ is nilpotent.*

Proof. Every nil two-sided ideal of any associative ring with minimal or maximal condition on left (or right) ideals is nilpotent; this is well known in the minimal case, and was proved by Levitzki [17] for the maximal case. Hence, by Theorem 4.1, R_1 is nilpotent, and so the Lie ring $[R]$ is clearly soluble. The nilpotence of $[R]$ now follows from Higgins' result (2).

We could, if we wished, strengthen the conclusion of Theorem 5.1 by asserting that R is even of finite class, i.e. that the lower central series $R=R_0 \geqq R_1 \geqq R_2 \geqq \cdots$ as defined by Jennings [13] terminates in zero; this follows from Jennings' Theorem 6.5. Also, by Gruenberg's result (1) and [7; Theorem 2], we see that every finitely generated associative Engel ring R with R_1 nilpotent has finite class. Now, if a given Engel ring R either has $k=k(y)$ or has prime characteristic or is π-regular, then we already know that R_1 is nil; hence, if R also satisfies the maximal condition on left (or right) ideals, then R_1 must be nilpotent, and our previous remark shows that R must have finite class locally. Again, by Theorem 4.5, every Engel ring with minimal condition on left (or right) ideals has finite class locally.

Having pointed out the most obvious consequences for Engel rings of our theorems of Section 4, we now round off this section with an account of some related results of Gruenberg (which have not previously appeared in print). Given any Lie ring L, any subset X of L and any non-negative integer r, let $X_{(r)}$ denote the set of all repeated commutators $[[\cdots[x_0,x_1],\cdots],x_r]$ with every $x_j \in X$.

THEOREM 5.2. *Let R be any finitely generated Engel ring, and X any given set of generators of R (as an associative ring). Then, if the Lie subring $[X]$ generated by X in $[R]$ is soluble, R must have finite class.*

Proof. We may clearly suppose that X is itself a finite set. Then $[X]$ is a finitely generated soluble Lie subring of an Engel ring, and is accordingly nilpotent by [8; Theorem 1′]: in particular, we can find a non-negative integer i such that $X_{(i)}$ contains only the zero element of R. If $i=0$ or 1 then of course $R_1=0$, while for $i\geqq 2$ [7; Theorem 2] gives

$$R_{(i-1)j}\leqq S_{(i)}+S_{(1)}{}^j \qquad (j=1,2,\cdots),$$

where $S_{(r)}$ denotes the two-sided ideal of R generated by $X_{(r)}$. Hence it is enough to prove that, for any $i\geqq 2$, $S_{(1)}{}^j\leqq S_{(i)}$ for sufficiently large j; by an obvious induction, we need only show that, for each positive integer r, there is a positive integer t such that $S_{(r)}{}^t\leqq S_{(r+1)}$.

To prove this, we observe first that, in any Engel ring, every central commutator $[a,b]$ is nilpotent. Indeed, if $[[a,b],ba]=0$, then we easily verify that $e_j(a,ba)=[a,b]^ja$ $(j=1,2,\cdots)$, whence $[a,b]^{k(a,ba)}a=0$, and similarly $[b,a]^{k(b,ab)}b=0$; letting $k=\max\{k(a,ba),k(b,ab)\}$, we at once conclude that

$$[a,b]^{k+1}=[a,b]^ka\cdot b-(-1)^k\cdot[b,a]^kb\cdot a=0.$$

Now define $R^*=R/S_{(r+1)}$ and, for any subset T of R, let T^* denote the image of T under the natural homomorphism of R onto R^*. Since X generates R, clearly $(X_{(r)})^*=(X^*)_{(r)}$ lies in the centre of R^*, and so, by applying the remark of the previous paragraph to the Engel ring R^*, every element of $(X^*)_{(r)}$ is nilpotent and central in R^*; consequently the ideal generated in R^* by the finite set $(X^*)_{(r)}$ is nilpotent, i.e. $S_{(r)}$ is nilpotent modulo $S_{(r+1)}$, as required.

Theorem 5.2 generalizes Jennings' result [14; Theorem 3] that, if a given finitely generated ring R has $[R]$ nilpotent, then R is of finite class. It also leads at once to

THEOREM 5.3. *If, in a given associative ring R, every finite subset X generates a nilpotent Lie ring $[X]$, then every finitely generated associative subring of R has finite class.*

To see this, we have only to apply Theorem 5.2 to the finitely generated subrings of R (noticing that the hypothesis of local Lie nilpotence always implies the Engel condition).

THEOREM 5.4. *If an Engel ring R has its index function $k(x,y)$ independent of both x and y, then R is locally of finite class.*

Proof. After Theorem 5.2, we have only to show that the Lie ring $[X]$ generated by any given finite subset X of R is soluble.

By Theorem 4.1, we know that R_1 is nil, while also, by our hypothesis on R, R_1 satisfies a polynomial identity $e_k(x,y) \equiv 0$ for some fixed k; hence, by Kaplansky [16; Theorem 5], R_1 is locally nilpotent (in the associative sense). Since X is finite we know, by [8; Theorem 2′], that the derived ring $[X]^2$ of $[X]$ is finitely generated as a Lie ring, say by the finite set Y. Then Y, being a finite subset of the locally nilpotent ring R_1, generates a nilpotent associative subring of R_1; and, since this subring of course contains $[X]^2$, clearly the Lie ring $[X]$ is soluble, as required.

We remark that the example of Baer [3] already mentioned in Section 2 shows that not every Engel ring with zero characteristic and $k = k(y)$ is locally of finite class. It remains open whether every Engel ring R with zero characteristic and $k(x,y)$ independent of both x and y has $[R]$ nilpotent (or, equivalently, soluble).

6. Weak K-rings. To extend our K-ring theorems of Section 4 to more general types of ring, we turn first to generalizations within the class of m-rings; since every division ring is an m-ring for all m, we exclude some of them by requiring in the defining property that $a = 0$, $a = 0$, i.e. we confine attention to those m-rings ($m = 1, 2, \cdots$) to each choice of x, y in which there correspond integers k, l, n such that $1 \leqq l \leqq m$, $x^{m-l} e_k(-x^l, y^n) = 0$. We shall in fact consider only the case $l = m$, but it turns out that we can usefully generalize this specialization of the m-rings by allowing m to depend on x, y: formally, we call a given ring R a *weak K-ring* whenever, to each choice of x, y in R, there correspond positive integers $k = k(x,y)$, $m = m(x,y)$ and $n = n(x,y)$ such that

$$e_k(x^m, y^n) = 0.$$

Obviously every K-ring is a weak K-ring. We note also that certain weak K-rings which do not at first sight seem to be K-rings nevertheless turn out to be on closer inspection:

THEOREM 6.1. *Let R be a weak K-ring of characteristic zero, whose index functions $k(x,y)$, $m(x,y)$, $n(x,y)$ are either* (a) *all independent of x, or* (b) *all independent of y. Then R is a K-ring with respectively either* (a) $e_{k(x,y)}(x, y^{m(x,y)n(x,y)}) = 0$ *or* (b) $e_{k(y,x)}(x, y^{m(y,x)n(y,x)}) = 0$ *for all x, y; in particular, by Theorem 4.1, R is commutator-nil.*

Proof. (a) Take any y in R (fixed throughout), and write for brevity

$$k(x,y) = k,\ m(x,y) = m,\ n(x,y) = n,\ y^n = z,$$

so that, for any $x, u \in R$, $e_k((x+hu)^m, z) = 0$ $(h = 0, 1, \cdots, m-1)$. Since

R has characteristic zero we deduce by the familiar "linearization argument" (cf. for example Higgins [11]) that

$$e_k(xu^{m-1} + uxu^{m-2} + \cdots + u^{m-1}x, z) = 0.$$

If R has a unit element 1, then, on taking $u = 1$ in this, we obtain $e_k(mx, z) = 0$, i.e. $e_k(x, z) = 0$, whence it is easy to see that $e_k(x, z^m) = 0$; we have to show that this conclusion holds independently of the presence of a unit.

Now clearly $[xz^{m-1} + zxz^{m-2} + \cdots + z^{m-1}x, z] = [x, z^m]$, and so, taking $u = z$, we have

$$\begin{aligned} e_{k-1}([x, z^m], z) &= e_{k-1}([xz^{m-1} + \cdots + z^{m-1}x, z], z) \\ &= e_k(xz^{m-1} + \cdots + z^{m-1}x, z) = 0. \end{aligned}$$

Also we can easily verify (by induction) that, for any $b \in R$,

$$e_j(b, z^m) = \sum_{i_1=1}^{m} \cdots \sum_{i_j=1}^{m} z^{i_1+\cdots+i_j-j} e_j(b, z) z^{jm-i_1-\cdots-i_j} \qquad (j = 1, 2, \cdots);$$

in particular, $e_j(b, z^m)$ vanishes whenever $e_j(b, z)$ does. Hence, taking $j = k-1$, $b = [x, z^m]$, we find $e_k(x, z^m) = e_{k-1}([x, z^m], z^m) = 0$, as required.

(b) Take any y in R (fixed throughout), and write now

$$k(y, x) = k,\ m(y, x) = m,\ n(y, x) = n,\ y^m = z,$$

so that, for any $x, u \in R$, $e_k(z, (x + hu)^n) = 0$ $(h = 0, 1, \cdots, kn-1)$. Linearizing once more and singling out the terms linear in x, we find

$$\sum_{i=0}^{k-1} e_{k-i-1}([e_i(z, u^n), xu^{n-1} + \cdots + u^{n-1}x], u^n) = 0.$$

On putting $u = z$ in this, each summand with $i \neq 0$ clearly reduces to 0; hence the summand corresponding to $i = 0$ must also vanish under the substitution $u = z$, i.e. we have

$$\begin{aligned} 0 &= e_{k-1}([e_0(z, z^n), xz^{n-1} + \cdots + z^{n-1}x], z^n) \\ &= e_{k-1}([z, xz^{n-1} + \cdots + z^{n-1}x], z^n) = e_{k-1}([z^n, x], z^n) \\ &= -e_{k-1}([x, z^n], z^n) = -e_k(x, y^{mn}), \end{aligned}$$

as required.

Theorem 6.1 (b), together with Lemma 4.2 and Theorem 4.3, enables us to make the following addition to our collection of K-ring results:

THEOREM 6.2. *Every K-ring whose index functions k, n depend only on x is commutator-nil.*

To show that every weak K-ring with k, m, n depending only on x (or only on y) is commutator-nil, it would, by Lemma 4.2, be enough to deal with the cases of zero and prime characteristic. Theorem 6.1 settles the first case, while in the second case the argument of Lemma 4.1 (ii) shows that we may take $k=1$; thus it would even suffice to prove that every ring with prime characteristic and $[x^{m(x)}, y^{n(x)}]=0$ is necessarily commutator-nil. But the writer has not yet been able to settle this critical case.

Our next two results extend Theorem 2.2, and the main part of the argument of Theorem 4.4, to all weak K-rings.

THEOREM 6.3. *Every primitive weak K-ring is a division ring.*

Proof. This is trivial (and still along essentially the same lines as that of Theorem 2.2). We may take our ring to be a dense set M of (right) linear transformations of a vector space V over a division ring D, and have only to show that V cannot contain two D-independent elements v_0, v_1. Now, if such v_0, v_1 existed, we could, since M is dense on V, find x, y in M such that $v_0x=v_1$, $v_1x=v_1$, $v_0y=0$, $v_1y=v_1$; but then, for any k, m, n,

$$v_0e_k(x^m, y^n)=v_0x^my^{kn}=v_1y^{kn}=v_1\neq 0$$

(since v_0, v_1 are D-independent), which is impossible if M is a weak K-ring.

THEOREM 6.4. *If a division ring D has the weak K-property, i.e. if there correspond to each pair x, y of elements of D integers $k(x,y)$, $m(x,y)$, $n(x,y)$ such that $e_{k(x,y)}(x^{m(x,y)}, y^{n(x,y)})=0$, then in fact $[x^{m(x,y)}, y^{n(x,y)}]=0$ for all x, y, i.e. we may always replace $k(x,y)$ by 1 while keeping $m(x,y)$ and $n(x,y)$ as given.*

Proof. Since D is a division ring, it has characteristic either a prime or zero; we consider these two cases separately.

If D has prime characteristic p, then, given any $x, y \in D$, and writing (for this half of the proof only) $m(x,y)=m$, $n(x,y)=n$, we can find a positive integer $r=r(x,y)$ such that $p^r \geqq k(x,y)$; then, by the argument of Lemma 4.1 (ii), we have

$$[x^m, (y^n)^{p^r}]=e_{p^r}(x^m, y^n)=0.$$

If $x\neq 0$, then x^{-m} exists, and

$$\begin{aligned} 0 &= x^{-m}[x^m, (y^n)^{p^r}] = (y^n)^{p^r} - x^{-m}(y^n)^{p^r}x^m \\ &= (y^n)^{p^r} - (x^{-m}y^nx^m)^{p^r} = (y^n - x^{-m}y^nx^m)^{p^r}; \end{aligned}$$

hence, since D is a division ring,

$$0 = y^n - x^{-m}y^n x^m = x^{-m}[x^m, y^n],$$

i.e. $[x^{m(x,y)}, y^{n(x,y)}] = 0$.

Suppose on the other hand that D has zero characteristic. If x, y existed in D with $[x^{m(x,y)}, y^{n(x,y)}] \neq 0$, then, writing $y^{n(x,y)} = z$, we could (as in the proof of Theorem 4.4) find a positive integer j such that $c = e_j(x^{m(x,y)}, z) \neq 0$, $[c, z] = 0$; on writing $d = -c^{-1}e_{j-1}(x^{m(x,y)}, z)$, it would follow from this that $[z, d] = 1$. Again, writing $zd = f$, we proceed to deduce a contradiction from $[z, f] = z \neq 0$. We now start from the relations

$$[z^s, f^q] = \{(f+s)^q - f^q\}z^s \qquad (q, s = 1, 2, \cdots),$$

which can be verified by induction on q as before (though the case $q = 1$ is no longer vacuous); it follows at once that

$$e_i(z^s, (f+hs)^q) = \{(f+(h+1)s)^q - (f+hs)^q\}^i z^s$$
$$(h, i = 0, 1, \cdots; q, s = 1, 2, \cdots).$$

Since D is a K-ring, $e_k(z^m, f^n) = 0$ for suitable positive integers k, m, n, and so

$$(f+hm)^n - f^n = 0 \qquad (h = 1, 2, \cdots, n),$$

whence, with $\Delta = 1!2!\cdots n!$ as before,

$$\Delta C_0{}^n m^n = \Delta C_1{}^n m^{n-1} f = \cdots = \Delta C_{n-1}{}^n m f^{n-1} = 0,$$

our desired contradiction.

In view of Theorems 6.3 and 6.4, to prove *every* weak K-ring commutative modulo its Jacobson radical, it would suffice to show that every division ring D subject to $[x^{m(x,y)}, y^{n(x,y)}] = 0$ (for all $x, y \in D$) is commutative. This is yet another unsolved problem; indeed, it is not even known whether commutativity holds when D has prime characteristic and m, n are either (a) both independent of x, or equivalently (b) both independent of y (which would at any rate imply that every weak K-ring R with m, n satisfying (a) or (b) has R/J commutative). However, by Theorem 4.2, we can at least assert

THEOREM 6.5. *Let R be any ring over which are defined integral-valued functions $k(x, y)$, $m(x, y)$ such that $e_{k(x,y)}(x^{m(x,y)}, y) = 0$ for all $x, y \in R$. Then R is commutative modulo its Jacobson radical.*

TRINITY COLLEGE, CAMBRIDGE.

REFERENCES.

[1] A. S. Amitsur, *Journal of the London Mathematical Society*, to appear.

[2] R. F. Arens and I. Kaplansky, "Topological representation of algebras," *Transactions of the American Mathematical Society*, vol. 63 (1948), pp. 457-481.

[3] R. Baer, "Radical ideals," *American Journal of Mathematics*, vol. 65 (1943), pp. 537-568.

[4] P. M. Cohn, "A non-nilpotent Lie ring satisfying the Engel condition and a non-nilpotent Engel group," *Proceedings of the Cambridge Philosophical Society*, vol. 51 (1955), pp. 401-405.

[5] M. P. Drazin, "Rings with central idempotent or nilpotent elements," to appear.

[6] ———, "A generalization of polynomial identities in rings," to appear.

[7] M. P. Drazin and K. W. Gruenberg, "Commutators in associative rings," *Proceedings of the Cambridge Philosophical Society*, vol. 49 (1953), pp. 590-594.

[8] K. W. Gruenberg, "Two theorems on Engel groups," *Proceedings of the Cambridge Philosophical Society*, vol. 49 (1953), pp. 377-380.

[9] I. N. Herstein, "A thorem on rings," *Canadian Journal of Mathematics*, vol. 5 (1953), 238-241.

[10] ———, "The structure of a certain class of rings," *American Journal of Mathematics*, vol. 75 (1953), pp. 864-871.

[11] P. J. Higgins, "Lie rings satisfying the Engel condition," *Proceedings of the Cambridge Philosophical Society*, vol. 50 (1954), pp. 8-15.

[12] N. Jacobson, *The Theory of Rings* (American Mathematical Society, 1943).

[13] S. A. Jennings, "Central chains of ideals in an associative ring," *Duke Mathematical Journal*, vol. 9 (1942), pp. 341-355.

[14] ———, "On rings whose associated Lie rings are nilpotent," *Bulletin of the American Mathematical Society*, vol. 53 (1947), pp. 593-597.

[15] I. Kaplansky, "A theorem on division rings," *Canadian Journal of Mathematics*, vol. 3 (1951), pp. 290-292.

[16] ———, "Rings with a polynomial identity," *Bulletin of the American Mathematical Society*, vol. 54 (1948), pp. 575-580.

[17] J. Levitzki, "Solution of a problem of G. Koethe," *American Journal of Mathematics*, vol. 67 (1945), pp. 437-442.

[18] M. Nagata, "On the nilpotency of nil-algebras," *Journal of the Mathematical Society of Japan*, vol. 4 (1952), pp. 296-301.

[19] M. Zorn, "On a theorem of Engel," *Bulletin of the American Mathematical Society*, vol. 43 (1937), pp. 401-404.

A CHARACTERIZATION OF PLANE CONTINUOUS CURVES WHICH ARE FILLED UP BY CONTINUOUS COLLECTIONS OF CONTINUOUS CURVES.*

By MARY-ELIZABETH HAMSTROM.

1. Introduction and notation. Let M denote a compact continuous curve in the plane. The purpose of this paper is to establish necessary and sufficient conditions on M that it should be the sum of the elements of a continuous collection G of mutually exclusive nondegenerate continuous curves. It is well known that if such a collection does fill up M then G is a continuous curve with respect to its elements and it is proved in [5] that G is an arc or a simple closed curve. Theorem 3 of [3] states that there is a continuous collection of arcs which fills up M and is an arc with respect to its elements if and only if M is a closed disc (a simple closed curve plus its interior) and Theorem II of [1] states that there is a continuous collection of simple closed curves filling up M if and only if M is an annulus. (In this connection, see also [2].)

This paper will be restricted to the consideration of points in the plane, E. Throughout this paper the letter M will be used to denote a compact continuous curve in E, $B(M)$ will denote the boundary of M and $U(M)$ will denote the collection of all boundaries of complementary domains of M. The notation $T(M)$ will denote the collection of all nondegenerate subcontinua x of M such that x is maximal with respect to the property of being a continuum which is the limiting set of an infinite sequence of distinct elements of $U(M)$. The notation $S(M)$ will denote the collection consisting of the elements of $U(M) + T(M)$ and all points[1] of $B(M) - [U^*(M) + T^*(M)]$.

Definition 1. If S is an upper semicontinuous collection of continua in the plane and if S, with respect to its elements, is an arc with endelements a and b then an arc h which is a non-endelement of S is said to be a K-arc or K-curve of S provided there exist two simple closed curves, J_1 and J_2,

* Received May 12, 1954; revised June 23, 1955. Presented to the Society, December 30, 1953.

[1] If K is a collection of point sets, K^* will denote the sum of the elements of K.

with mutually exclusive interiors, I_1 and I_2, such that $J_1 \cdot J_2$ is an arc which contains and is separated by h, $J_1 \cdot J_2 \cdot S^* = h$, I_1 does not intersect any element of the interval[2] ah of elements of S and I_2 does not intersect any element of the interval hb. If h is a or b then h is a K-curve provided there is a simple closed curve containing h but intersecting no other element of S. A simple closed curve which is an element of S is said to be a K-curve of S provided it either separates two points of S^* in E or is an endelement of S. A point which is an element of S is said to be a K-curve of S provided it is not the sequential limiting set of any sequence of simple closed curves in S.

Definition 2. If V is a closed subcollection of arcs of S then if n is a natural number $V^{(n)}$ will denote the collection of all elements g of V such that if ϵ is a positive number and U is an open subcollection of S containing g then U contains an element of V which has a subarc the distance between whose endpoints is less than ϵ but whose diameter is not less than $1/n$.

Definition 3. A collection S of mutually exclusive point sets is said to be imbedded in a collection W of mutually exclusive point sets provided that each element of S is a subset of some element of W and each element of W contains one and only one element of S. (These definitions were used in [6].)

2. Statement of the principal theorem, some examples and some properties of M.

THEOREM 2. *The compact plane continuous curve M is the sum of the elements of a continuous collection of mutually exclusive nondegenerate continuous curves which is an arc with respect to its elements if and only if every boundary of a complementary domain of M is a simple closed curve and there is an upper semicontinuous collection S such that* (1) *S is an arc with respect to its elements,* (2) *$B(M)$ is a subset of S^*,* (3) *every non-endelement of S is either an element of $T(M)$, a point of $M - B(M)$, a point of $B(M) - [U^*(M) + T^*(M)]$ or a point or pair of points of some element of $U(M)$,* (4) *each endelement of S is a nondegenerate continuous curve which does not separate M and is a continuum of condensation*[3] *of M,* (5) *each non-endelement of S which is an element of $T(M)$ is an arc or*

[2] If H is an upper semicontinuous collection of point sets which is an arc with respect to its elements and x and y are elements of H then the interval xy of elements of H is the collection consisting of x, y, and all elements of H between x and y. The segment xy is the interval xy minus the elements x and y.

[3] A nondegenerate subcontinuum K of M is a continuum of condensation of M provided each point of K is a limit point of $M - K$.

simple closed curve which is a K-curve of S and does not intersect any element of $U(M)$ in more than one point and (6) *if H is a closed subcollection of arcs which are elements of S then for no natural number n is it true that* $H^{(n)} = H$.

Theorem 2 gives a condition on the boundary of M. Example 4 below, although fairly elementary, shows that condition (4) in the statement of Theorem 2 permits the boundary of M to have some rather complicated components and so long as these components can be made subsets of the endelements of S the conditions (1) through (6) can be satisfied. This prevents the formulation of a less complicated characterization.

Suppose V is an upper semicontinuous collection of mutually exclusive arcs, simple closed curves and points which fills up a compact plane continuum and is an arc with respect to its elements. In [6] (Theorem 2) I proved that there exists (in the plane) a continuous collection W of simple closed curves such that V is imbedded in W if and only if (1) each element of V is a K-curve of V, (2) if H is a closed subcollection of arcs of V then for no natural number n is it true that $H^{(n)} = H$ and (3) of each pair of simple closed curves which are elements of V one is interior to the other. This theorem will be applied in a certain way to the elements of S in order to show that the conditions of the statement of Theorem 2 are sufficient.

Example 1. Suppose M consists of two concentric circles C_1 and C_2 plus the points between them. Let t_1 and t_2 be mutually exclusive arcs lying in $M - (C_1 + C_2)$. Let P_1 and Q_1 and P_2 and Q_2 be pairs of diametrically opposite points of C_1 and C_2 respectively. Let s_1 and s_2 be mutually exclusive arcs each lying except for its endpoints in $M - (C_1 + C_2 + t_1 + t_2)$ where one endpoint of s_i is P_i, the other a point of t_i $(i = 1, 2)$. Let s be an arc lying except for its endpoints Q_1 and Q_2 in $M - (C_1 + C_2 + t_1 + t_2 + s_1 + s_2)$. Then if S is the collection consisting of t_1, t_2, the points of $s_1 + s_2 + s$ not in $t_1 + t_2$, and the pairs of points which are the endpoints of the chords of C_i perpendicular to the diameter P_iQ_i $(i = 1, 2)$, the conditions (1) through (6) of Theorem 2 are satisfied. If C_2 is interior to C_1 then the collection G may consist of (1) t_1 and t_2, (2) a simple closed curve g_1 having $C_2 + s_2 + t_2 + s$ in its exterior and $t_1 + s_1 - P_1$ in its interior and such that $g_1 \cdot C_1 = g_1 \cdot s_1 = P_1$, (3) a simple closed curve g_2 having $g_1 + t_1 + s_1$ in its exterior and $C_2 + s_2 + t_2 + s - Q_1$ in its interior and such that $g_2 \cdot C_1 = Q_1$, (4) a simple closed curve g_3 having $g_2 + s - Q_2$ in its exterior and $t_2 + s_2 + C_2 - Q_2$ in its interior and such that $g_3 \cdot C_2 = Q_2$, (5) a simple closed curve g_4 having $g_3 + C_2 - P_2$ in its exterior and $t_2 + s_2 - P_2$ in its

interior and such that $g_4 \cdot C_2 = P_2$, (6) appropriate simple closed curves each lying interior to g_1 or g_4, each having t_1 or t_2 in its interior and each having only one point in common with s_1 or s_2, (7) appropriate arcs each either lying exterior to both g_1 and g_2 and having for endpoints the endpoints of a chord of C_1 perpendicular to P_1Q_1 or lying interior to g_3 and exterior to g_4 and having for endpoints the endpoints of a chord of C_2 perpendicular to P_2Q_2, and (8) appropriate simple closed curves lying interior to g_2 and exterior to g_3 and having only one point in common with s.

In the more general cases the construction of G will be patterned after the above construction, which is a natural way to proceed.

Example 2. Let M' be a simple closed curve J plus its interior I, let K denote a contracting sequence of mutually exclusive closed discs such that K^* is a subset of I which is dense in I and let M denote M' minus the sum of the interiors of the elements of K. Clearly there is no collection satisfying the conditions (1) through (6) in the statement of Theorem 2.

Example 3. Let t be an arc lying in the interior I of a closed disc M', let α be a contracting sequence of mutually exclusive closed discs such that t is the limiting set of α and α^* lies in $I - t$ and let M denote M' minus the sum of the interiors of the elements of α. If we choose t as one of the elements of $T(M)$ it is seen that in this case a collection S satisfying conditions (1) through (6) does exist.

Example 4. Let J be a simple closed curve lying in the interior of a closed disc M' and let J_1, J_2, and J_3 be mutually exterior simple closed curves each lying in the interior of M', each having one and only one point in common with J and each having its interior lying in the exterior of J. Let M denote M' minus the sum of the interiors of J, J_1, J_2, and J_3. It is seen that if $J_1 + J_2 + J_3 + J$ is an endelement of S then a collection satisfying the conditions of Theorem 2 does exist. There is a continuous collection of continuous curves filling up M, one endelement of which is $J + J_1 + J_2 + J_3$, the other endelement is the outer boundary of M and the non-endelements are simple closed curves whose interiors contain $J + J_1 + J_2 + J_3$.

Example 5. It has been suggested by a referee that condition (6) of Theorem 2 can be omitted. The author even announced at the April, 1955 meeting of the Society in Brooklyn, N. Y. that if V is an upper semicontinuous collection of arcs and points such that (1) V is an arc with respect to its elements and fills up a compact plane continuum and (2) each non-degenerate element of V is a K-curve of V and a continuum of condensation

of V^* then if H is a closed subcollection of arcs of V then for no natural number n is it true that $H^{(n)} = H$. If this were a true statement condition (6) would follow immediately from conditions (1) through (5) and the above mentioned Theorem 2 of [6]. However, the following example shows that this statement is not true.

Let T be a Cantor set in the interval $(0,1)$ of the X axis. If x is a point of $(0, 1)$ let T_x be a vertical straight line interval from the point $(x,0)$ to the point $(x,1)$. Let $I_1, I_2, \cdots$ be the components (which are segments) of $(0,1) - T$, let $x_1, x_2, \cdots$ and $z_1, z_2, \cdots$ be the right and left hand endpoints of these segments and let $y_1, y_2, \cdots$ be their midpoints. For each j let S_{x_j} be an arc which is the sum of T_{x_j}, T_{y_j} and the straight line interval from the point $(y_j, 1)$ to the point $(x_j, 0)$. Let H be the collection consisting of S_{x_j} for each j and T_x for each x in $T - \Sigma x_j$. For each j let G_j be an upper semicontinuous collection which is an arc with respect to its elements such that (1) each non-endelement of G_j is a point, (2) the end-elements of G_j are S_{x_j} and T_{z_j}, (3) each point of $S_{x_j} - T_{x_j}$ is a limit point of $G_j^* - S_{x_j}$ but no point of $T_{x_j} - x_j$ is such a limit point and (4) every point of G_j^* lies on or between the lines $y = 0$ and $y = 1 + 1/i$ and on or between the lines $x = z_j$ and $x = x_j$. The collection consisting of the elements of ΣG_i and H is an upper semicontinuous collection V of arcs and points, it is an arc with respect to its elements and each of its nondegenerate elements is a K-curve of V and a continuum of condensation of V^*. However, $H^{(2)} = H$.

If J is a simple closed curve whose interior contains V^* and appropriate intervals of points of G_j are replaced by appropriate closed discs whose interiors are subsequently removed from the interior of J, we obtain an example of a continuous curve satisfying all the conditions of Theorem 2 except (6). A study of the construction of V in the above paragraph should convince the reader that this second example can be easily obtained.

Before proceeding to the proof of Theorem 2 I shall establish some lemmas. In what follows, the letter G will denote a continuous collection of nondegenerate continuous curves which fills up M and is an arc with respect to its elements.

LEMMA 1. *Every complementary domain of M is bounded by a simple closed curve and no two such boundaries have more than one point in common.*

Proof. It follows easily that M has no cut point. Theorem 54, Chapter IV of [8] states that if K is the boundary of a complementary domain D of a compact continuum N and A is a cut point of K then A is a cut point of N.

Consequently, no boundary of a complementary domain of M has a cut point and since M is locally connected every such boundary curve is a simple closed curve. Clearly no two such boundary curves have more than one point in common.

COROLLARY. *If the non-endelement g of G intersects the element J of $U(M)$, (the collection of all boundary simple closed curves of M), then $g \cdot J$ does not contain an arc.*

Proof. This corollary is a rather easy consequence of the fact that the elements of G and the closures of the components of $M - g$ are locally connected. A detailed proof may be found in Lemma 1 of [4].

LEMMA 2. *Each non-endelement of G is either an arc or a simple closed curve.*

Proof. Let g be a non-endelement of G and let U be an arc of elements of G one of whose endelements is g. The continuous collection U fills up the continuous curve U^*, hence, by Lemma 1, every boundary of a complementary domain of U^* is a simple closed curve. But since g is a limit element of $G - U$, g is a subset of such a boundary simple closed curve and is thus either an arc or a simple closed curve.

The following theorem follows directly from Theorem 8 of [4] but since I use the result quite frequently I shall prove it here in a modified form. It shows how the elements of G are related to each other and to the boundary simple closed curves of M.

THEOREM 1. *Let J be an element of $U(M)$ that is not a subset of either endelement of G and let K be the collection of all elements of G which intersect J. Then* (1) *each non-endelement of K is an arc g such that $g \cdot J$ is the sum of the endpoints of g,* (2) *no non-endelement of K intersects any element of $U(M)$ other than J,* (3) *each endelement h of K either is an endelement of G and $h \cdot J$ is connected or is a simple closed curve which is a non-endelement of G, does not intersect three elements of $U(M)$ and has only one point in common with J.*

Proof. The collection K is clearly an arc with respect to its elements. If the common part of J and the endelement h of K separates J then h separates two elements of K in M and consequently in K^*. This is impossible, so $h \cdot J$ is connected and if h is a non-endelement of G (and hence is an arc or a simple closed curve) it follows from the corollary to Lemma 1 that $h \cdot J$ is degenerate.

If g is a non-endelement of K, $g \cdot J$ separates J and is therefore not connected and, by the corollary to Lemma 1, $g \cdot J$ is totally disconnected. By Lemma 2, g is an arc or a simple closed curve. If $g \cdot J$ contains three points or if g is a simple closed curve, $M - g$ is the sum of three mutually separated point sets, which is not true. Hence g is an arc and it follows from the continuity of the collection G that the two points of $g \cdot J$ are the endpoints of g.

Suppose the above mentioned non-endelement g of K intersects an element J' of $U(M)$ different from J. Let K' denote the collection of all elements of G which intersect J'. (By the corollary to Lemma 1, K' is a nondegenerate collection.) Since g belongs to both K and K' and is a non-endelement of K there is an element g' of G which is a non-endelement of both K and K'. Since J and J' can not both contain the endpoints of g' a contradiction follows.

If an arc k lies in and separates M then k contains two points of some element of $U(M)$. For, if not, then if $M - k$ is the sum of the two mutually separated point sets V_1 and V_2 every element of $U(M)$ lies either in $V_1 + k$ or $V_2 + k$. Consequently $E - k$ is the sum of two mutually separated point sets U_1 and U_2 where U_i $(i = 1, 2)$ is V_i plus the sum of the complementary domains of M whose boundaries lie in $V_i + k$. Since no arc separates E this is impossible.

Consequently, if the endelement h of K is an arc and a non-endelement of G, h contain two points of an element J' of $U(M)$ and is a non-endelement of the collection of elements of G which intersect J'. But no such non-endelement intersects two boundary simple closed curves. Therefore h is a simple closed curve.

If the endelement h of K is a non-endelement of G (therefore a simple closed curve) and intersects three elements, J, J_1, and J_2 of $U(M)$, two of these elements—say J_1 and J_2—have the property that $J_1 - J_1 \cdot h$ and $J_2 - J_2 \cdot h$ lie in the same complementary domain of h. Since G is a continuous collection, some arc which is a non-endelement of G intersects both J_1 and J_2, which we have seen to be impossible. This completes the proof of Theorem 1.

It is to be noted that the interior of each simple closed curve which is a non-endelement of G contains an endelement of G. Thus the collection of non-endelements does not contain three mutually exterior simple closed curves. Also, if h and h' are non-endelements of G and are also the endelements of the collection K considered above, then the interiors of h, h' and J are not mutually exclusive. If this is so then each element of $G - K$ is either in

the interior of h or in the interior of h' and the non-endelements of K are exterior to both h and h'. However, the outer boundary of M is not interior to h or h' and is not J. Hence a non-endelement of K intersects both J and the outer boundary of M, which Theorem 1 has shown to be impossible. Hence the interiors of h, h' and J are not mutually exclusive. In particular, if J is not the outer boundary of M then one of h and h' is interior to the other.

LEMMA 3. *If the limiting set K of a sequence α of distinct elements of $U(M)$ is a nondegenerate continuum then K is a subset of some element of G.*[4]

Proof. Since M is locally connected and no two of its boundary simple closed curves have more than one point in common, every sequential limiting set of a subsequence of α is degenerate. Let G' denote the collection of all elements of G which intersect K. Let J be an element of α lying in G'^* but intersecting neither endelement of G and let V denote the collection of non-endelements of the collection of all elements of G intersecting J. The set V^* is open in M, but since no element of V intersects any element of $U(M)$ other than J, V^* does not intersect any element of $U(M)$ other than J. Since K intersects V^* and is the limiting set of α this is impossible. Hence K is a subset of some element of G.

LEMMA 4. *If the element J of $U(M)$ does not intersect either endelement of G then J does not contain three limit points of $B(M) - J$ (where $B(M)$ denotes the boundary of M).*[5]

Proof. Let H denote the collection of all elements of G which intersect J. It follows from the corollary to Lemma 1 that H is not degenerate. The sum, V, of the non-endelements of H is an open subset of M and it follows from Theorem 1 that V does not intersect any element of $U(M)$ different from J. Hence V does not intersect $B(M) - J$. Thus, if J contains limit points of $B(M) - J$ these points lie in the endelements of H. But by Theorem 1, the sum of the endelements of H has only two points in common with J. Thus Lemma 4 is proved.

LEMMA 5. *If a non-endelement u of G contains an element of $T(M)$, u is a simple closed curve.* (*See the definition of $T(M)$ in Section* 1.)

Proof. It follows from Lemma 2 that u is either an arc or a simple

[4] This also follows from Theorem 9 of [4].

[5] This is a slight extension of Lemma 2 of [4].

closed curve. If u is an arc the argument in the fourth paragraph of the proof of Theorem 1 proves that there is a boundary simple closed curve J of M such that $u \cdot J$ is not connected. Thus, by Theorem 1, u is a non-end-element of the collection H of all elements of G which intersect J. But by Theorem 1 again, V, the sum of the non-endelements of H, does not intersect any element of $U(M)$ other than J, which, since u contains the limiting set of an infinite sequence of distinct elements of $U(M)$, is impossible. Therefore u is a simple closed curve.

It is to be noted that it follows from Lemma 3 that every element of $T(M)$ is a subset of some element of G.

LEMMA 6. *If K_1 and K_2 are continua such that* (1) *$K_1 + K_2$ is a subset of $B(M)$ (the boundary of M) and does not intersect either endelment of G,* (2) *the common part of K_1 and K_2 is either a point or an element of $T(M)$ and* (3) *each of K_1 and K_2 contains an element of $U(M)$ and is the sum of some elements of $S(M)$, then $K_1 \cdot K_2$ does not contain a limit point of $B(M) - (K_1 + K_2)$.*

Proof. Suppose a point P of $K_1 \cdot K_2$ is a limit point of $B(M) - (K_1 + K_2)$ and let g denote the element of G which contains P. If P is a point of an element t of $T(M)$ then t is a subset of g, by Lemma 3. In any case, since $U^*(M)$ is a dense subset of $B(M)$, P is a limit point of the common part of $B(M) - (K_1 + K_2)$ and $U^*(M)$. Also, one of the two components of $M - g$—say L—has the property that P is a limit point of the common part of $B(M) - (K_1 + K_2)$, L and $U^*(M)$. Thus there is a sequence $J_1, J_2, \cdots$ such that P is a limit point of ΣJ_i and for each i, J_i belongs to $U(M)$ and intersects the common part of L and $B(M) - (K_1 + K_2)$.

Suppose one of the continua K_1 and K_2—say K_1—intersects L. There is a natural number n and a continuum U of elements of G such that every element of U is an arc which intersects J_n (and thus is a non-endelement of the collection of all elements of G which intersect J_n) and U^* separates some point of $K_1 \cdot L$ from $K_1 \cdot K_2$ in $\bar{L}$. It follows from the hypothesis to this lemma that some element of $U(M)$ intersects $K_1 \cdot U^*$, which, since U^* also intersects J_n is, by Theorem 1, impossible.

Therefore both K_1 and K_2 are subsets of g plus the other component, N, of $M - g$. There is an element J of $U(M)$ lying in K_1 and a subcontinuum U of elements of G such that every element of U is an arc which intersects J and U^* separates some point of $K_2 \cdot N$ from $K_1 \cdot K_2$ in $\bar{N}$. But K_2 contains an element of $U(M)$ which intersects U^*—which, again by Theorem 1, is impossible. Thus the supposition that $K_1 \cdot K_2$ contains a limit point of $B(M) - (K_1 + K_2)$ leads to a contradiction.

COROLLARY 1. *If K is a continuum which is a subset of $B(M)$ such that* (1) *the common part of K and an endelement h of G is either degenerate or an element of $T(M)$,* (2) *K contains an element of $U(M)$ which is not a subset of h and* (3) *K is the sum of some elements of $S(M)$ then h does not contain a limit point of $B(M) - (K + h)$.*

Proof. Since $M - h$ has only one component, the argument used in the proof of Lemma 6 can be used to prove that the assumption that h contains a limit point of $B(M) - (K + h)$ leads to a contradiction.

COROLLARY 2. *If two elements of $U(M)$ intersect h then at least one of them lies in h (where h is an endelement of G).*

Proof. If J is an element of $U(M)$ which intersects but is not contained in h, then h is an endelement of the (nondegenerate) collection of all elements of G which intersect J. But no non-endelement of this collection intersects any element of $U(M)$ other than J. Therefore, if an element J' of $U(M)$ other than J intersects h, J' is a subset of h.

LEMMA 7. *The collection T of all elements of G which contain elements of $T(M)$ is nowhere dense in G.*

Proof. Let V be an open subcollection of G. If an element v of V contains an element of $T(M)$ then V^* contains a boundary simple closed curve J of M. It follows from Theorem 1 and Lemma 5 that no non-endelement of the collection of all elements of V which intersect J contains any element of $T(M)$. The collection of all such non-endelements is an open subcollection of G. Thus every open subcollection of G contains an open subcollection of G which lies in $G - T$ and therefore T is nowhere dense in G.

It is to be noted that it is possible for an element of G to contain more than one element of $T(M)$.

3. Definition of a collection and proof that the conditions (1) through (6) of the statement of Theorem 2 are necessary. Let h and h' denote the endelements of G. If an element J of $U(M)$ intersects h or h'—say h—then $J \cdot h$ is connected, since h is an endelement of the collection H of all elements of G which intersect J. If an element t of $T(M)$ does not lie in (and consequently, by Lemma 3, does not intersect) $h + h'$, it follows from Lemmas 3 and 5 that t is a subset of a simple closed curve which is an element of G. Thus, by Theorem 1, if t intersects an element J of $U(M)$, $t \cdot J$ is degenerate.

If J is an element of $U(M)$ then (1) if J does not intersect $h + h'$

then J does not contain three limit points of $B(M)-J$, (2) if J intersects one of h and h'—say h—but not the other, $J-h\cdot J$ does not contain two limit points of $B(M)-J$, and (3) if J intersects both h and h', $J-J\cdot(h+h')$ does not contain any limit point of $B(M)-J$. Statement (1) follows from Lemma 4 and (2) and (3) follow from obvious modifications of the proof of Lemma 4.

If the element J of $U(M)$ intersects both h and h' let P_J and Q_J denote the continua $h\cdot J$ and $h'\cdot J$. If J intersects one of h and h'—say h—but not the other, let P_J denote the continuum $h\cdot J$ and if $J-h\cdot J$ contains a limit point of $B(M)-J$ (we have seen above that there are not two such points) denote it by Q_J; otherwise let Q_J denote any point of $J-h\cdot J$. If J does not intersect $h+h'$ and contains two limit points of $B(M)-J$ denote them by P_J and Q_J; if J contains only one limit point of $B(M)-J$ denote it by P_J and let Q_J denote any point of $J-P_J$; if J contains no limit point of $B(M)-J$, let P_J and Q_J denote any two points of J.

In any case, if J is not a subset of $h+h'$, $J-(P_J+Q_J)$ does not contain any limit points of $B(M)-J$ and, as was seen in the first paragraph of this section, if J intersects an element t of $T(M)$, $t\cdot J$ is a subset of P_J or Q_J. There is clearly an upper semicontinuous collection H_J of mutually exclusive point sets which fills up J, is an arc with respect to its elements and is such that its endelements are P_J and Q_J and each of its non-endelements is a pair of points separated in J by P_J+Q_J.

Let S' be the collection consisting of (1) h and h', (2) every element of $T(M)$ which does not intersect $h+h'$, (3) every element of H_J which does not lie in $h+h'+T^*(M)$, for all elements J of $U(M)$ and (4) the points of $B(M)-[h+h'+U^*(M)+T^*(M)]$. It follows from the construction of S' and the continuity of the collection G that S' is an upper semicontinuous collection.

LEMMA 8. *The collection S' contains no subcollection which is a simple closed curve with respect to its elements.*

Proof. Suppose K is a subcollection of S' which is a simple closed curve with respect to its elements. It follows from Lemma 6 and its corollaries that neither h nor h' is an element of K. Let J be an element of $U(M)$ which lies in K^* and let H be the collection of all elements of G which intersect J. The points P_J and Q_J do not lie in the same endelement of H, since neither endelement of H has two points in common with J. But, by Theorem 1, the sum of the non-endelements of H does not intersect any element of $U(M)$ other than J and hence does not intersect K^*-J, since

$U^*(M)$ is dense in $B(M)$ and therefore in S'^*. Therefore, since P_J and Q_J lie in different endelements of H the collection H_J of all elements of S' lying in J separates two elements of K in K, these two elements of K lying in different components of $M - H^*$. Since H_J is an arc with respect to its elements, this is impossible. Hence the supposition that a subcollection K of S' is a simple closed curve with respect to its elements leads to a contradiction.

It now follows from the lemmas, particularly Lemma 4 and Lemma 6 and its corollaries, that S' contains no subcollection that is a triod with respect to its elements, that every component of S' is, with respect to its elements, either an arc or a point and that if the nondegenerate component K of S' contains a limit element of $S' - K$, that limit element is an endelement of K.

Let W be the collection whose elements are the elements of S' and the points of $M - S'^*$. Clearly W is an upper semicontinuous collection filling up M, each nondegenerate cyclic element of W (the elements of W regarded as points) is homeomorphic to a two-sphere, every arc t of elements of W has a non-endelement which is a limit element of $W - t$ and no ordinary element or cyclic element of W separates W into three mutually separated subcollections. From the construction of W it is seen that every separating element of W is a simple closed curve and consequently an element of $T(M)$. Moreover, if (1) z is an arc of elements of S', considered as a subcollection of W, (2) s is an element of z, (3) z lies in a two-sphere of elements of W, and (4) s is a separating element of W then s is an endelement of z.

In a paper of R. L. Moore and J. R. Kline (see [7]) it is proved that, in the plane, in order for a closed, compact point set K to be a subset of an arc it is necessary and sufficient that every component of K should be a point or an arc t and no point of t, except, possibly, for the endpoints of t, is a limit point of $K - t$. This theorem clearly holds on a two-sphere and it is seen from the construction of the collection W and the remarks of the above paragraph that in W, S' satisfies the conditions required of K in the Moore-Kline Theorem so that this theorem can be applied to the space of elements of W. Let S be an upper semicontinuous subcollection of W such that S contains S' and S, with respect to its elements, is an arc with endelements h and h'. The collection S is the collection required by Theorem 2.

It follows directly from the construction of S that S satisfies conditions (1), (2), and (3). Condition (4) follows from the fact that h and h' are the endelements of the continuous collection G and condition (5) follows from Lemmas 3 and 5, Theorem 1 and the corollary to Lemma 1. That is, if the

non-endelement t of S is an element of $T(M)$, it is, by Lemma 3, a subset of a non-endelement of G and hence is an arc or a simple closed curve. It follows from Lemma 5 that t is a subset of a simple closed curve which is an element of G. From this and the continuity of G it follows that t is a K-curve of S. The other requirements of (5) follow from Theorem 1 and the corollary to Lemma 1.

A special case of a theorem of J. H. Roberts ([9], Theorem 2) states that if E is a continuous collection of arcs filling up a compact metric space then there is a subcollection E' of E such that E' is dense in E and E is equicontinuous[6] at each element of E'. An argument suggested by that outlined in [9] yields a similar result in the case where E is a continuous collection of arcs and simple closed curves. Since each closed subcollection of arcs of S is imbedded in a continuous subcollection of arcs and simple closed curves of G then it follows from Roberts' result that if H is such a closed subcollection then for no natural number n is it true that $H^{(n)} = H$.

4. Proof that the conditions are sufficient. Suppose that M is a compact plane continuous curve and that S is an upper semicontinuous collection of subsets of M such that M and S satisfy the conditions required of them in the statement of Theorem 2. Since the subcollection of S consisting of the elements of $T(M)$ is a nowhere dense subcollection it is seen that M is the closure of a domain.

Let h and h' denote the endelements of S. It follows from conditions (1) to (3) that if $h + h'$ intersects an element of $T(M)$ then $h + h'$ contains that element and if one of h and h'—say h—intersects an element J of $U(M)$ then $J \cdot h$ is connected. Let T denote the collection of elements of $T(M)$ which lie in $M - (h + h')$. From the remarks of the above paragraph it follows that T is a subcollection of S whose closure is totally disconnected in S. It also follows from conditions (1), (2), (3), and (5) that if an element t of T intersects an element J of $U(M)$ then $t \cdot J$ is degenerate.

If H_J denotes the collection of all elements of S which intersect the element J of $U(M)$ and J is not a subset of $h + h'$ then H_J is an arc with respect to its elements. Let k and k' denote the endelements of H_J. Clearly $k \cdot J$ and $k' \cdot J$ are connected. Each of k and k' is either h or h', an element

[6] The collection E of mutually exclusive arcs and simple closed curves is said to be equicontinuous at an element h of E provided that for any $\epsilon > 0$ and any point p of h there exists a $\delta > 0$ such that if x and y are points of the same element h' of E and are each within δ of p then there is an arc in h' containing $x + y$ and of diameter less than ϵ.

of T or a point of J. Each non-endelement of H_J is a pair of points of J which separates $J \cdot k$ from $J \cdot k'$ in J.

Let K denote a two-sphere. There is a continuous transformation f carrying M onto K such that for each point x of K, $f^{-1}(x)$ is either (1) h or h', (2) an element of $S - (T + h + h')$, (3) a point of T^* or (4) a point of $M - [h + h' + U^*(M)]$. (I.e. We "sew" up the boundary simple closed curves of M.) Let $f(T)$ denote the collection of all point sets $f(t)$ where t is an element of T; let $f(S)$ denote the collection of all point sets $f(S)$ where s is an element of S.

Clearly the collection $f(S)$ is an upper semicontinuous collection of atriodic continuous curves and is an arc with respect to its elements. From the definition of K-curve for a simple closed curve and property (5) of the statement of Theorem 2 it follows that on K, if interiors be taken with respect to one of the endelements of $f(S)$, the elements of $f(S)$ which are simple closed curves have the property that of any two of them one is interior to the other. Therefore, from property (6) and a very slight modification [7] of the proof of the above mentioned Theorem 2 of [6], (see the second paragraph following the statement of Theorem 2), it follows that there is a continuous collection W filling up K such that $f(S)$ is imbedded in W, the endelements of W are the endelements of $f(S)$ and the non-endelements of W are simple closed curves. If G denotes the collection of all continua $f^{-1}(w)$ where w is an element of W, G is the collection required by Theorem 2, since, by condition (4) and the construction of f, G is continuous and each of its elements is a nondegenerate continuous curve. This completes the proof of Theorem 2.

5. Collections which are simple closed curves with respect to their elements.

THEOREM 3. *The compact plane continuous curve M is the sum of the elements of a continuous collection of nondegenerate continuous curves which is a simple closed curve with respect to its elements if and only if M is an annulus.*

Proof. The condition is obviously sufficient. To prove that it is necessary let G denote a continuous collection of nondegenerate continuous curves which fills up M and is a simple closed curve with respect to its elements and let g be an element of G. It follows from an argument in the proof of Lemma 2 that g is either an arc or a simple closed curve and it follows from

[7] Theorem 2 of [6] is proved for point sets in the plane, but the proof is easily carried over to the two-sphere.

the fact that g separates G locally but does not separate the continuous curve M that g is an arc. From Theorem 3 of [3] (stated in the first paragraph of this paper) it follows that G^* is the sum of two closed discs (simple closed curves plus their interiors) whose common part is the sum of two mutually exclusive arcs. Thus M is an annulus.

GOUCHER COLLEGE.

REFERENCES.

[1] R. D. Anderson, "Continuous collections of continuous curves in the plane," *Proceedings of the American Mathematical Society*, vol. 3 (1952), pp. 647-657.

[2] Herman Cohen, "Some results concerning homogeneous plane continua," *Duke Mathematical Journal*, vol. 18 (1951), pp. 467-474.

[3] Mary-Elizabeth Hamstrom, "Concerning continuous collections of continuous curves," *Proceedings of the American Mathematical Society*, vol. 4 (1953), pp. 240-243.

[4] ———, "Concerning webs in the plane," *Transactions of the American Mathematical Society*, vol. 74 (1953), pp. 500-513.

[5] R. D. Anderson and Mary-Elizabeth Hamstrom, "A note on continuous collections of continuous curves filling up a continuous curve in the plane," *Proceedings of the American Mathematical Society*, vol. 5 (1954), pp. 748-752.

[6] Mary-Elizabeth Hamstrom, "Concerning the imbedding of upper semicontinuous collections of continua in continuous collections of continua," *American Journal of Mathematics*, vol. 76 (1954), pp. 793-810.

[7] R. L. Moore and J. R. Kline, "On the most general plane closed point set through which it is possible to pass a simple continuous arc," *Annals of Mathematics*, vol. 20 (1919), pp. 219-223.

[8] R. L. Moore, "Foundations of point set theory," *American Mathematical Society Colloquium Publications*, vol. 13 (1932).

[9] J. H. Roberts, "Collections filling a plane," *Duke Mathematical Journal*, vol. 2 (1936), pp. 10-19.

COMMUTATOR SUBGROUPS OF FREE GROUPS.*[1]

By MAURICE AUSLANDER and R. C. LYNDON.

Let F be a non-abelian free group, R a non-trivial normal subgroup of F, and $G = F/R$. Then F and R are free groups of ranks at least two. Let $[R, R]$ be the commutator subgroup of R, and let $F_0 = F/[R, R]$ and $R_0 = R/[R, R]$. Then R_0 is a free abelian group of rank at least two. Moreover, the inner automorphisms of F induce a natural operation of G on R_0 which we denote by $g \cdot r_0$.

THEOREM 1. *The operation of G on R_0 is effective; that is, only the unit element of G leaves all elements of R_0 fixed.*

Proof. We may suppose $G \neq 1$. Let $g \neq 1$ in G generate a cyclic group G', and let F' be the inverse image of G' under the canonical map of F onto G. Then $G' = F'/R$, and the operation of G' on R_0 agrees with that of G on R_0. We must show that g, and hence that G', does not operate trivially on R_0.

Since $G' = F'/R$ where F' is a free group, the cup product reduction theorem [2, Theorem 10.1] asserts that for any G'-module A

$$H^3(G', A) \simeq H^1(G', \mathrm{Hom}(R_0, A)), \tag{1}$$

where the operation of G' on $\mathrm{Hom}(R_0, A)$ is defined by

$$(g \cdot \theta)(r_0) = g \cdot [\theta(g^{-1} \cdot r_0)], \quad \theta \in \mathrm{Hom}((R_0, A), \quad r_0 \in R_0.$$

We take A to be isomorphic with the cyclic group $G' = Z_n$, $n \geqq 0$, with G' operating trivially on A. Suppose that G' operates trivially on R_0 as well. Then all operations in (1) are trivial. Since G' is cyclic, $H^3(G', A)$ is cyclic or trivial [2, 16.2]. On the other hand $H^1(G', \mathrm{Hom}(R_0, A))$ is isomorphic to $\mathrm{Hom}(Z_n, \mathrm{Hom}(R_0, Z_n))$ and hence to $\mathrm{Hom}(Z_n \otimes R_0, Z_n)$ where $Z_n \otimes R_0$ denotes the tensor product of Z_n and R_0. Since R_0 is a free abelian group of rank at least two, $Z_n \otimes R_0$ is a direct product of at least two copies of Z_n.

* Received February 25, 1955; revised May 16, 1955.

[1] Part of the work presented in this paper was done while the first named author was at the University of Chicago. Substantial simplification of the proofs, and a strengthening of Theorem 3, are due to R. L. Taylor.

But $Z_n \neq 1$. Therefore, $\mathrm{Hom}(Z_n \otimes R_0, Z_n)$ is neither cyclic nor trivial. This contradicts (1). Therefore G' does not operate trivially on R_0.

For the following corollaries, we asume that $R \neq F$.

COROLLARY 1.1. *$[R,R]$ is a proper subgroup of $[F,R]$.*

Proof. Let $f \in F$, $f \notin R$. If $[f,r] = frf^{-1}r^{-1}$ were in $[R,R]$ for all r in R, then, setting $g = fR$ in G, we should have $g \cdot r_0 = r_0$ for all $r_0 = r[R,R]$ in R_0. This contradicts Theorem 1.

COROLLARY 1.2. *$[R,R]$ is a proper subgroup of $[F,F]$.*

COROLLARY 1.3. *The center C_0 of F_0 is a proper subgroup of R_0.*

Proof. We first show that $C_0 \subset R_0$. Suppose f_0 is in C_0. Then $f_0 r_0 = r_0 f_0$ for all r_0 in R_0. This implies that the image $g = f_0 R$ of f_0 in G operates trivially on R_0. By Theorem 1, this implies that $g = 1$. Hence f_0 is in R_0.

Suppose that $C_0 = R_0$. Then $f_0 r_0 = r_0 f_0$ for all f_0 in F_0 and r_0 in R_0. Hence all g in G operate trivially on R_0. By Theorem 1, this is possible only if $G = 1$, which contradicts the hypothesis that $R \neq F$.

COROLLARY 1.4. *The center C_0 of F_0 consists precisely of those elements of R_0 that are left fixed under the operation of G.*

THEOREM 2. *The center C_0 of F_0 is trivial if and only if G is infinite.*

Proof. Let G be of finite order n, with elements $g_1 = f_1 R, \cdots, g_n = f_n R$. Suppose $C_0 = 1$. For any r in R, and corresponding $r_0 = r[R,R]$ in R_0, the element $\prod_{i=1}^{n} g_i \cdot r_0$ is fixed under G, hence in C_0. Therefore $\prod g_i \cdot r_0 = 1$. This implies that $\prod f_i r f_i^{-1} [R,R] \subset [F,F]$. Since $f_i r f_i^{-1} \equiv r \bmod [F,F]$, this implies that $r^n \in [F,F]$, where $n \geqq 1$. Since $F/[F,F]$ is torsion free, r is in $[F,F]$. Thus $C_0 = 1$ implies $R \subset [F,F]$. But this would imply that the free abelian group $F/[F,F]$ were a homomorphic image of the finite group $G = F/R$, a contradiction. Therefore $C_0 \neq 1$.

Let G be infinite, with elements $g_0 = 1, g_1, g_2, \cdots$. Let ZG be the integer group ring of G. Its elements are of the form $\gamma = \sum n_i g_i$, where only a finite number of the integers n_i are not zero. Let G operate on ZG by left multiplication. If γ were fixed under G, then $g_i^{-1} \cdot \gamma = \gamma$ for all i. This implies that all $n_i = n_0$, and hence that all n_i are zero. Then $\gamma = 0$. It follows that in any free G-module only the zero element is fixed under G. Since [cf. 1, Ch. XIV, Ex. 11] R_0 is a submodule of a free G-module, R_0 is without non-trivial fixed elements. Therefore, by Corollary 1.4, C_0 is trivial.

For any group G, let $G_2 = [G, G]$, $G_3 = [G_2, G]$.

LEMMA. *If F is a non-abelian free group, and R a normal subgroup such that $F_2 \subset R$ and $F_3 \subset R_2$, then $F = R$.*

Proof. From $F_3 \subset R_2$ it follows that $F_2/R_2 \subset C_0$. If $R \neq F$, then $R_2 \neq F_2$ by Corollary 1.2, whence $C_0 \neq 1$, and, by Theorem 2, $G = F/R$ is finite. Thus, in any case, G is finite.

The commutator operation determines a homomorphism

$$\kappa\colon F/F_2 \otimes F/F_2 \to F_2/F_3.$$

This induces a homomorphism $\mu\colon G \otimes R_0 \to R_0$; explicitly, $\mu(g \circ r_0) = g \cdot r_0 - r_0$. Since G is finite, $G \otimes R_0$ is a torsion group. But R_0 is a free abelian group. It follows that $\mu = 0$. This means that $R_0 \subset C_0$. By Corollary 1.3, this implies that $R = F$.

THEOREM 3. *If S and R are normal subgroups of a non-abelian free group, and $[S, S] \subset [R, R]$, then $S \subset R$.*

Proof. Let $F = RS$. We may suppose that F is a non-abelian free group. We have

$$F_2 = [F, F] = [RS, RS] = [R, R][R, S][S, S] \subset R_2(R \cap S),$$

$$F_3 = [F_2, F] \subset [R_2(R \cap S), RS] \subset R_2[R, R][S, S] = R_2.$$

The lemma now gives $R = F$, whence $S \subset R$.

UNIVERSITY OF MICHIGAN.

REFERENCES.

[1] H. Cartan and S. Eilenberg, Homological Algebra, Princeton University Press, 1955.

[2] S. Eilenberg and S. MacLane, "Cohomology theory in abstract groups. I," *Annals of Mathematics*, vol. 48 (1947), pp. 51-78.

ERRATA.

P. Hartman and A. Wintner, "Asymptotic integrations of linear differential equations," this JOURNAL, vol. 77 (1955), pp. 45-86 and 404.

On page 47, formula (xii), read $h_* - h_0$ in place of $h_0 - h_*$ in the exponent of t.

On page 60, line 13, read (13) in place of (12).

On page 62, formula (67), and on page 72, formula (119), read $\epsilon(a)$ in place of the factor $(\epsilon(j) + \epsilon(a) - 1)$ in the exponent of t.

On page 63, formula (70), read $>$ instead of $<$ (twice); in place of the two lines following (70), read "In this case, $y(t)$ is unique."

On page 64, line 11, read [6] in place of [5].

On page 64, third line of Section 12, read $y \to e^{-\mu t}y$ in place of $y \to e^{-\mu t}$.

On page 66, formula (80), read $(b(q) - k)$ in place of the factor $(b(q) - h)$.

On page 67, at the end of formula (87), read $+k_0$ in place of $-k_0$.

On page 68, formula (91), read $(\epsilon(a) - \epsilon(j))$ in place of the factor $(\epsilon(a) - 1)$ in the exponent of t.

On page 69, in the first line of formula (97), read $k > b$ in place of $k < b$.

On page 83, formula (171), read 1 in place of -1.

On page 83, formulae (172) and (174), read $(1 - f'^2/16f^3)$ in place of $(1 + f'^2/16f^3)$.

On page 84, formula (175), read $(1 - a)$ in place of $(1 + a)$.

On page 84, formula (177), read $(f^{\frac{1}{2}}(t) - a(t))$ in place of $(f^{\frac{1}{2}}(t) + a(t))$.

On page 84, beginning of formula (182), read $(f^{-\frac{1}{2}}z')'$ in place of $(f^{-\frac{1}{2}}z')$.

ON THE PREPARATION OF MANUSCRIPTS.

The following instructions are suggested or dictated by the necessities of the technical production of the American Journal of Mathematics. Authors are urged to comply with these instructions, which have been prepared in their interests.

Manuscripts not complying with the standards usually have to be returned to the authors for typographic explanations or revisions and the resulting delay often necessitates the deferment of the publication of the paper to a later issue of the Journal.

Horizontal fraction signs should be avoided. Instead of them use either solidus signs / or negative exponents.

Neither a solidus nor a negative exponent is needed in the symbols $\frac{1}{2}$, $\frac{1}{\pi}$, $\frac{1}{2\pi}$, $\frac{1}{2\pi i}$, which are available in regular size type.

Binomial coefficients should be denoted by C_k^n and not by parentheses. Correspondingly, for symbols of the type of a quadratic residue character the use of some non-vertical arrangement is usually imperative.

Use the exponent $\frac{1}{2}$ instead of the square root sign.

Replace $e(\ )$ by $\exp(\ )$ if the expression in the parenthesis is complicated.

By an appropriate choice of notations, avoid unnecessary displays.

Simple formulae, such as $A + iB = \frac{1}{2}C^*$ or $s_n = a_1 + \cdots + a_n$, should not be displayed (unless they need a formula number).

Use $'$ or d/dx, possibly D, but preferably not a dot, in order to denote ordinary differentiation and, as far as possible, a subscript in order to denote partial differentiation (when the symbol ∂ cannot be avoided, it should be used as $\partial/\partial x$).

Commas between indices are usually superfluous and should be avoided if possible.

In a determinant use a notation which reduces it to the form $\det a_{ik}$.

Subscripts and superscripts cannot be printed in the same vertical column, hence the manuscript should be clear on whether a_i^k or $a^k{}_i$ is preferred. (Correspondingly, the limits of summation must not be typed *after* the Σ-sign, unless either $\sum_1^m$ or $\sum^m{}_1$ is desired.) If a letter carrying a subscript has a prime, indicate whether x_i' or x'_i is desired.

Experience shows that a tilde or anything else *over* a letter is very unsatisfactory. Such symbols often drop out of the type after proof-reading and, when they do not, they usually appear uneven in print. For these reasons we advise against their use. This advice applies also to a bar over a Greek or German letter (for the symbol of complex conjugation an asterisk is often allowed by the context). Type carrying bars over ordinary size italic letters of the Latin alphabet is available.

Bars reaching over several letters should in any case be avoided (in particular, type lim sup and lim inf instead of lim with upper and lower bars).

Repeated subscripts and superscripts should be used only when they cannot be avoided, since the index of the principal index usually appears about as large as the principal index. Bars and other devices *over* indices cannot be supplied. On the other hand, an asterisk or a prime (to be printed *after* the subscript) is possible on a subscript. The same holds true for superscripts.

Distinguish carefully between l. c. "oh," cap. "oh" and zero. One way of distinguishing them is by underlining one or two of them in different colors and explaining the meaning of the colors.

Distinguish between ϵ (epsilon) and ε (symbol), between x (eks) and $\times$ (multiplication sign), between l. c. and cap. phi, between l. c. and cap. psi, between l. c. k and kappa and between "ell" and "one" (for the latter, use $\mathfrak{l}$ and l respectively).

Avoid unnecessary footnotes. For instance, references can be incorporated into the text (parenthetically, when necessary) by quoting the number in the bibliographic list, which appears at the end of the paper. Thus: "[3], pp. 261-266."

Except when informality in referring to papers or books is called for by the context, the following form is preferred:

[3] O. K. Blank, "Zur Theorie des Untermengenraumes der abstrakten Leermenge," *Bulletin de la Societé Philharmonique de Zanzibar*, vol. 26 (1891), pp. 242-270.

In any case, the references should be precise, unambiguous and intelligible.

Usually section numbers and section titles are printed in bold face, the titles "Theorem," "Lemma" and "Corollary" are in caps and small caps, "Proof," "Remark" and "Definition" are in italics. This (or a corresponding preference) should be marked in the manuscript. Use a period, and *not* a colon, after the titles Theorem, Lemma, etc.

German, script and bold face letters should be underlined in various colors and the meaning of the colors should be explained. The same device is needed for Greek letters if there is a chance of ambiguity. In general, mark all cap. Greek letters.

All instructions and explanations for the printer can conveniently be collected on a separate sheet, to be attached to the manuscript.

In case of doubt, recent issues of the Journal may be consulted.

THE JOHNS HOPKINS PRESS · · BALTIMORE 18

American Journal of Hygiene. Edited by D. BODIAN, Managing Editor, A. M. BAETJER, R. M. HERRIOTT, M. MERRELL, P. E. SARTWELL, and T. B. TURNER. Publishing two volumes of three numbers each year, volume 62 is now in progress. Subscription $12 per year. (Foreign postage, 50 cents; Canadian postage, 25 cents.)

American Journal of Mathematics. Edited by R. BAER, W.-L. CHOW, A. WEIL and A. WINTNER. Quarterly. 8 vo. Volume 77 in progress. $8.50 per volume. (Foreign postage, 50 cents; Canadian, 25 cents.)

American Journal of Philology. Edited by H. T. ROWELL, LUDWIG EDELSTEIN, K. MALONE, B. D. MERITT, JAMES H. OLIVER, and EVELYN H. CLIFT, Secretary. Volume 76 is in progress. $6.00 per volume. (Foreign postage, 50 cents; Canadian 25 cents.)

Bulletin of the History of Medicine. OWSEI TEMKIN, Editor. Bi-monthly. Volume 29 in progress. 8vo. Subscription $6 per year. (Foreign postage, 50 cents; Canadian 25 cents.)

Bulletin of the Johns Hopkins Hospital. Edited by PHILIP F. WAGLEY. Monthly. 8vo. Subscription $8 per year. Vol. 96-97 is in progress. (Foreign postage, 50 cents; Canadian, 25 cents.)

ELH. A Journal of English Literary History. Edited by E. T. NORRIS (Managing Editor) and D. C. ALLEN, G. E. BENTLEY, F. MICHAEL KROUSE, C. P. LYONS, J. HILLIS MILLER, ROY H. PEARCE and E. R. WASSERMAN. Quarterly 8vo. Volume 22 in progress. $4.00 per volume. (Foreign postage, 40 cents; Canadian, 20 cents.)

Johns Hopkins Studies in Romance Literatures and Languages. †H. C. LANCASTER, Editor. 8vo. Seventy-six numbers have been published.

Johns Hopkins University Studies in Archaeology. DAVID M. ROBINSON, Editor. 8vo. Thirty-nine volumes have appeared.

Johns Hopkins University Studies in Geology. JOSEPH T. SINGEWALD, JR., Editor. 8vo. Sixteen numbers have been published.

Johns Hopkins University Studies in Historical and Political Science. Under the direction of the Departments of History, Political Economy and Political Science. 8vo. Volume 72 in progress.

Modern Language Notes. Edited by KEMP MALONE, DON CAMERON ALLEN, CHARLES R. ANDERSON, ANNA HATCHER, WILLIAM H. MCCLAIN, GEORGES POULET, BRUCE W. WARDROPPER, EARL R. WASSERMAN, and S. K. HENINGER, JR., *Secretary*. Eight times yearly. 8vo. Volume 70 in progress. $6 per volume. Foreign postage, 50 cents; Canadian, 25 cents.)

A complete list of publications will be sent on request

Lectures given at the Johns Hopkins University
(spring term, 1953) by

CARL LUDWIG SIEGEL

TOPICS IN CELESTIAL MECHANICS

Mimeographed. Bound. ii + 237 pages. Baltimore 1955.

Price $3.50 postpaid. Orders should be sent and checks made payable to

THE JOHNS HOPKINS BOOKSTORE, BALTIMORE 18, MD.